Limit Theorems for Slowly Mixing Dynamical Systems

by

Matt Galton

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Mathematics Institute

September 2019

Contents

List of Figures

Acknowledgments

Firstly, huge thanks go to my supervisor Ian Melbourne for his constant support and guidance throughout my PhD. I couldn't have asked for a better supervisor.

Thanks must also go to my family, for their support and serving as a welcome respite throughout my PhD. To my dad, thanks for the uncountable curries and beer. To my mum, thanks for lending a kind ear. To my brother, thanks for distracting me from my PhD with rugby and comedy nights when needed, and paying for all of our lunches!

To Cerys, words can't describe how much your optimism, love and support has helped me over the last four years. Sneaking doodles into my notes never failed to make me laugh. We've made Marie proud! Thanks also to my second family, Stephen, Ann, Becky, Tom and Jess, for the laughs and letting me use your home for holiday trips in the sun.

To everyone that shared a coffee or an evening playing computer games with me, pretending to listen to my problems, thanks.

Finally, thanks to Warwick's Maths department for the wine and cheese, and the Stats department for the coffee!

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree, apart from the material in Chapter 2 and Sections 3.1, 3.2, which was previously submitted to the University of Warwick for the degree of Masters of Science. This material has undergone significant change, and new results have been added since submission.

I plan on publishing a paper on the non-invertible and invertible iterated weak invariance principle.

Abstract

In this thesis we prove the iterated weak invariance principle for ergodic, probability-preserving dynamical systems with respect to L^∞ observables under a mild mixing assumption. When the dynamics can be modelled by a Young tower the iterated weak invariance principle is already known under optimal conditions. The setting where T is not necessarily modelled by a Young tower still has gaps, however. This is the setup considered in this thesis, and it is flexible enough to include time-one maps of suitable (semi-)flows. For both non-invertible and invertible maps, we improve upon the previous best results. For non-invertible maps, our mixing assumption is optimal when correlations decay at polynomial rates.

Chapter 1

Introduction

Many chaotic dynamical systems have similar properties to random processes, such as the law of large numbers (or Birkhoff's ergodic theorem), the central limit theorem (CLT) (see for example [9, 13, 22, 28, 32, 47, 50]) and the weak invariance principle (WIP) (see [9, 13, 22, 50]).

For limit theorems in continuous time, a standard mechanism exists whereby it often suffices to show a limit theorem for a suitable Poincaré map [39]. This has been successfully applied to billiard flows [3], geodesic flows over negatively curved manifolds [13, 47] and the classical Lorenz equations [23], amongst others. Consequently this thesis only considers discrete time dynamical systems.

Many advances have focussed on particular dynamical systems, such as the collision map for various billiards [9], expanding maps of the interval with a neutral fixed point [33, 45, 53] and Hénon maps [4]. In this thesis we take an abstract approach: Let $T : \Lambda \to \Lambda$ be an ergodic, measure-preserving transformation over a probability space $(\Lambda, \mathcal{A}, \mu)$ and let $v : \Lambda \to \mathbb{R}^d$ be a mean zero, $L^\infty(\Lambda, \mathbb{R}^d)$ observable. From an abstract perspective, the literature available for limit theorems in discrete time is vast [5, 11, 14, 17, 25, 31, 34, 35] and various different techniques are available [18, 43, 49]. We focus on the martingale approach of Gordin [18]. The main idea is to relate the observable v to a martingale through a coboundary and apply readily available martingale limit theorems [21].

For $t \in [0, 1]$ define $W_n(t) := n^{-1/2} \sum_{j=0}^{[nt]} v \circ T^j$, and define the $d \times d$ matrix-valued process $\mathbb{W}_n^{ij}(t) := \int_0^t W_n^i \, dW_n^j$, where the integral is understood in the Stieltjes sense. We say $(W_n, \mathbb{W}_n)$ satisfies an *iterated weak invariance principle* (iterated WIP) if $(W_n, \mathbb{W}_n) \to_w (W, \mathbb{W})$ holds over the Skorohod space $\big(D([0, 1], \mathbb{R}^d \times \mathbb{R}^{d\times d}), \|\cdot\|_\infty\big)$, where $\to_w$ denotes weak convergence, W is a suitable d dimensional Brownian motion and $\mathbb{W}$ is a $d \times d$ matrix-valued process.

Motivating the Iterated WIP

Note that W_n actually depends on two variables, $W_n(x,t) = n^{-1/2}\sum_{j=0}^{[nt]-1} v(T^j x)$. Roughly speaking, if we choose $x \in \Lambda$ randomly according to μ, then we can consider $W_n(x,\cdot) : [0,1] \to \mathbb{R}^d$ as a random path. The WIP says that on very large time scales, W_n looks like random noise. More precisely $W_n \to_w W$.

A natural question is to suppose that W_n are driving some differential equations, and ask if the solutions converge in some sense. This is known as homogenization [10, 19, 29, 30, 38]. Since W_n converges to a random noise, one might hope that the solutions converge to a process containing a random noise component. To be more precise, define $z_n(t) = [nt]/n$ and suppose $a \in C^{1+}(\mathbb{R}^m, \mathbb{R}^m)$ and $b \in C^{2+}(\mathbb{R}^m, \mathbb{R}^{m\times d})$. Let x_n be the unique m-dimensional solution of the (discontinuous) ODE

$$dx_n = a(x_n^-)dz_n + b(x_n^-)dW_n, \quad x_n(0) = \zeta \in \mathbb{R}^m, \tag{1.1}$$

where ζ is fixed. We hope that x_n converges weakly to some random process X, and X is the unique solution to an SDE of the form

$$dX = a(X)dt + b(X) \,\square\, dW, \quad X(0) = \zeta, \tag{1.2}$$

where $b(X) \,\square\, dW$ is a suitable stochastic integral (see [10, Theorem 4.10] for more details). Unfortunately, the WIP alone does not uniquely determine the weak limit of the iterated integrals $\int W_n^i \, dW_n^j$, so things can become problematic:

Example. In equation (1.1) take $m = d = 2$, $a \equiv 0$, $b(x_n^1, x_n^2) = \begin{pmatrix} 1 & 0 \\ 0 & x_n^1 \end{pmatrix}$ and $\zeta = 0$. We now have

$$dx_n^1 = dW_n^1, \quad dx_n^2 = x_n^1 dW_n^2.$$

Therefore $x_n^1(t) = W_n^1(t)$ and $x_n^2(t) = \int_0^t W_n^1 dW_n^2$. The mapping $\chi : D([0,1], \mathbb{R}^2) \to D([0,1], \mathbb{R}^2)$, $\chi(g)(t) = (g^1(t), \int_0^t g^1 \, dg^2)$ is not continuous and so the WIP $W_n \to_w W$ does not guarantee weak convergence of x_n.

The saving grace is that understanding the weak limit of the iterated integrals is nearly enough to guarantee convergence of x_n:

Theorem ([10, Theorem 4.10])**.** *If* $(W_n, \mathbb{W}_n)$ *satisfies an iterated WIP and a suitable moment bound (see [10, Assumption 4.7]), then* $x_n \to_w X$*, where* x_n *and* X *are solutions to equations* (1.1) *and* (1.2) *respectively.*

In this thesis we focus on the iterated WIP for both invertible and non-invertible maps. For ease of exposition, the following discussion considers correlations decaying at polynomial rates.

Non-Invertible maps

Suppose for every $n \geq 1$, $i = 1, \ldots, d$ and $w \in L^\infty(\Lambda)$ we have

$$\left| \int v^i \, w \circ T^n \, d\mu \right| \leq C \|w\|_\infty n^{-\alpha},$$

where $C > 0$ and $\alpha > 0$.

The CLT and WIP hold when $\alpha > 1$ and are due to Liverani [32, Theorem 1.1] and Dedecker and Rio [11, Theorem 1], respectively.

If T is modelled by a Young tower [52, 53], then the iterated WIP is known under the optimal condition $\alpha > 1$ [29, Theorem 10.2]. By optimal we mean that there are examples where the CLT fails for $\alpha \leq 1$ [20].

When T is not necessarily modelled by a Young tower, the iterated WIP holds under the hypothesis $\alpha > 2$ by [29, Theorem 4.3]. In this thesis, we improve this result to $\alpha > 1$, so the iterated WIP holds exactly when the CLT holds. Our result is optimal for correlations decaying at polynomial rates. We also give examples where this new result applies.

Invertible maps

The mixing condition for invertible maps is slightly different to that of non-invertible maps: Assume there exists a sub-sigma algebra $\mathcal{A}_0 \subset \mathcal{A}$ such that $T^{-1}\mathcal{A}_0 \subset \mathcal{A}_0$ and assume that for every $n \geq 1$,

$$\|E_0(v \circ T^{-n})\|_1 \leq C n^{-\alpha}, \quad \|E_0(v \circ T^n) - v \circ T^n\|_1 \leq C n^{-\beta},$$

where $C > 0$ and $\alpha, \beta > 0$. The first condition is analogous to the mixing condition in the non-invertible case.

The CLT and WIP hold when $\alpha, \beta > 1$ and are due to Liverani [32, Theorem 1.2] and Dedecker and Rio [11, Corollary 4(b)], respectively.

If T is modelled by a Young tower, take $\mathcal{A}_0$ as the sigma algebra generated by the stable foliation. Then the first condition corresponds to the rate of decay of correlations for the quotient map, and the second condition corresponds to the average rate of contraction along stable leaves [12]. The iterated WIP is known under the condition $\alpha, \beta > 1$ [40, Corollary 2.3].

When T is not necessarily modelled by a Young tower, the iterated WIP holds for $\alpha, \beta > 2$ [29, Theorem 5.2]. In this thesis we improve this to $\alpha > 1, \beta > 2$. This result may have room for further improvement as the CLT/WIP holds when $\alpha, \beta > 1$. We give an example not covered by previous work.

Thesis layout

In Chapter 2 we overview some basic preliminary material on mixing.

Chapter 3 contains the limit theorems for non-invertible dynamical systems: Section 3.2 proves the CLT and recovers part of a well known result by Liverani [32, Theorem 1.1]. We use a different decomposition due to Melbourne [36, Equation (2.1)], yielding a more elementary proof. Section 3.3 proves the WIP using this decomposition, setting up some preliminary calculations needed for the iterated WIP. Section 3.4 proves the iterated WIP, and gives an example which is not covered by previous known work.

In Chapter 4 we prove the corresponding invertible limit theorems and provide an example not covered by previous work.

We write $a_n \ll b_n$ if there is a constant $C > 0$ such that $a_n \leq C b_n$ for all $n \geq 1$. The letter C is reserved for *some* positive constant.

Chapter 2

Preliminary Material

2.1 Rates of Mixing

Let $(\Lambda, \mathcal{A}, \mu)$ be a probability space and $T : \Lambda \to \Lambda$ be a measurable map. We say T is *measure-preserving* with respect to μ if for every $A \in \mathcal{A}$ we have $\mu(T^{-1}A) = \mu(A)$. An equivalent definition of measure-preserving is that $\int v \circ T \, d\mu = \int v \, d\mu$ for every $v \in L^1(\Lambda)$. We say T is *ergodic* with respect to μ if for all $A \in \mathcal{A}$ such that $T^{-1}A = A$ we have $\mu(A) = 0$ or 1.

Let $v : \Lambda \to \mathbb{R}$ be a measurable function. We call v an *observable*. The observable v can be thought of as taking 'observations' of a system as it evolves, $(v(x), v(Tx), v(T^2x), \ldots)$. Let

$$v_n := \sum_{j=0}^{n-1} v \circ T^j$$

be the n^{th} *Birkhoff sum* of v (a subscript n after any observable denotes the n^{th} Birkhoff sum).

Proposition 2.1.1 (Birkhoff's Ergodic Theorem [44, Theorem 2.3])**.** *Let $T : \Lambda \to \Lambda$ be an ergodic, measure-preserving transformation over the probability space $(\Lambda, \mathcal{A}, \mu)$, and let $v \in L^1(\Lambda)$. Then*

$$\frac{1}{n}\sum_{j=0}^{n-1} v \circ T^j \to_{a.s.} \int v \, d\mu.$$

□

Suppose T is a measure-preserving transformation with respect to μ. One method for measuring how unpredictable the dynamics of a system are is to see how well correlated a 'measurement at time zero' $v(x)$ and a 'measurement at time n' $v(T^n x)$ are. We measure this using the *correlation function*

$$\mathrm{C}_n(v, w) := \int v\, w \circ T^n \, d\mu - \int v \, d\mu \int w \, d\mu.$$

The map T is called *mixing* with respect to μ if for every $v, w \in L^2(\Lambda)$,

$$C_n(v, w) \to 0$$

as $n \to \infty$.

Mixing can be arbitrarily slow with respect to $L^\infty(\Lambda)$ observables, so we restrict ourselves to consider observables in smaller Banach spaces $\mathcal{L}$, such as bounded variation or Hölder observables. We say that $T : \Lambda \to \Lambda$ *mixes exponentially* with respect to $\mathcal{L}$ if there exists constants $K > 0,\ c > 0$ and a suitable test space $\mathcal{K}$ such that $\mathrm{C}_n(v, w) \leq K e^{-cn}$ for every $n \geq 1$, $v \in \mathcal{L}$ and $w \in \mathcal{K}$. We say that $T : \Lambda \to \Lambda$ *mixes polynomially* with respect to $\mathcal{L}$ if there exists constants $K > 0,\ \beta > 0$ and a suitable test space $\mathcal{K}$ such that $\mathrm{C}_n(v, w) \leq K n^{-\beta}$ for every $n \geq 1$, $v \in \mathcal{L}$ and $w \in \mathcal{K}$. The rate of mixing indicates how quickly a system 'forgets' its initial condition.

Let $p \in [1, \infty]$ and $q^{-1} = 1 - p^{-1}$ be its conjugate, with $q = 1$ when $p = \infty$. Define the *Koopman operator*

$$U : L^p(\Lambda) \to L^p(\Lambda), \quad Uv := v \circ T.$$

The Koopman operator enjoys the following properties:

- $\int Uv \, d\mu = \int v \, d\mu$.
- $\|Uv\|_p = \|v\|_p$ for all $v \in L^p(\Lambda),\ p \in [1, \infty]$.
- $\int Uv\, Uw \, d\mu = \int v\, w \, d\mu$ for all $v \in L^p(\Lambda),\ w \in L^q(\Lambda)$.

The *transfer operator* $P : L^q(\Lambda) \to L^q(\Lambda)$ is defined as the unique operator satisfying

$$\int v\, w \circ T \, d\mu = \int Pv\, w \, d\mu \text{ for all } v \in L^q(\Lambda),\ w \in L^p(\Lambda).$$

The transfer operator has the following properties:

- $P1 = 1$.
- $\int Pv \, d\mu = \int v \, d\mu$.
- $\|Pv\|_p \leq \|v\|_p$ for all $p \in [1, \infty]$.

- $PU = I$.

Proposition 2.1.2. *Let $T : \Lambda \to \Lambda$ be a measure-preserving map. If T is invertible, then $UP = I$ and $\|Pv\|_p = \|v\|_p$ for all $v \in L^p(\Lambda)$ and $p \geq 1$. In general we have $UP = \mathbb{E}(\,\cdot\,|\,T^{-1}\mathcal{A})$.*

Proof. Let $v \in L^1(\Lambda)$ and $w \in L^\infty(\Lambda)$. If T is invertible, then

$$\int Pv\, w\, d\mu = \int v\, w \circ T\, d\mu = \int v \circ T^{-1} \circ T\, w \circ T\, d\mu = \int v \circ T^{-1}\, w\, d\mu.$$

Therefore $Pv = v \circ T^{-1}$ and $UP = I$. The fact $\|Pv\|_p = \|v\|_p$ for any $p \geq 1$ now follows easily from invariance of T.

Next suppose T is a general measure-preserving map and let $A \in \mathcal{A}$. By uniqueness of conditional expectation we must show $\int_{T^{-1}A} UPv\, d\mu = \int_{T^{-1}A} v\, d\mu$ and that UPv is $T^{-1}\mathcal{A}$-measurable. The first condition follows from the following calculation:

$$\int_{T^{-1}A} UPv\, d\mu = \int_\Lambda U\mathbb{1}_A\, UPv\, d\mu = \int_\Lambda \mathbb{1}_A\, Pv\, d\mu = \int_\Lambda U\mathbb{1}_A\, v\, d\mu = \int_{T^{-1}A} v\, d\mu.$$

Next, since $P : L^1(\Lambda) \to L^1(\Lambda)$, we see Pv is $\mathcal{A}$-measurable and hence UPv is $T^{-1}\mathcal{A}$-measurable. $\square$

For any $n \geq 1$, the transfer operator for T^n is P^n, since

$$\int v\, w \circ T^n\, d\mu = \int Pv\, w \circ T^{n-1}\, d\mu = \cdots = \int P^n v\, w\, d\mu.$$

Proposition 2.1.3. *Suppose $T : \Lambda \to \Lambda$ is a measure-preserving map and let $\gamma_n \geq 0$, for $n \geq 1$. Then for any $v \in L^1(\Lambda)$ with $\int v\, d\mu = 0$ the following statements are equivalent:*

- $\|P^n v\|_1 \leq \gamma_n$,
- *For every $w \in L^\infty(\Lambda)$ we have $|C_n(v, w)| \leq \|w\|_\infty \gamma_n$.*

Proof. First suppose $\|P^n v\|_1 \leq \gamma_n$. Then for any $w \in L^\infty(\Lambda)$ we have

$$|C_n(v, w)| = \left|\int v\, w \circ T^n\, d\mu\right| = \left|\int P^n v\, w\, d\mu\right| \leq \|P^n v\|_1 \|w\|_\infty \leq \|w\|_\infty \gamma_n.$$

Next suppose that $|C_n(v, w)| \leq \|w\|_\infty \gamma_n$ for any $w \in L^\infty(\Lambda)$. Then,

$$\|P^n v\|_1 = \left|\int P^n v\, \operatorname{sign}(P^n v)\, d\mu\right| \leq \|\operatorname{sign}(P^n v)\|_\infty \gamma_n = \gamma_n.$$

$\square$

Rates of mixing depend on the regularity of the observable, as we see from the two following examples.

Example 2.1.4. Let $T : [0,1] \to [0,1]$ be the doubling map, $Tx = 2x \mod 1$, and let $v : [0,1] \to \mathbb{R}$ be a Lipschitz continuous observable. For any $n \geq 1$ we have

$$\Big\| P^n v - \int v \, d\mu \Big\|_\infty \leq |v|_{\text{Lip}} \, 2^{-n},$$

where $|v|_{\text{Lip}} := \sup\limits_{x \neq y} \dfrac{|v(x) - v(y)|}{|x-y|}$ is the Lipschitz constant of v.

To see this, first note that $(Pv)(x) = \dfrac{1}{2}\left[v\left(\dfrac{x}{2}\right) + v\left(\dfrac{x+1}{2}\right)\right]$, so

$$\begin{aligned}
|(Pv)(x) - (Pv)(y)| &= \left| \frac{1}{2}\left[v\left(\frac{x}{2}\right) - v\left(\frac{y}{2}\right)\right] + \frac{1}{2}\left[v\left(\frac{x+1}{2}\right) - v\left(\frac{y+1}{2}\right)\right]\right| \\
&\leq \frac{1}{2}|v|_{\text{Lip}} \left| \frac{x}{2} - \frac{y}{2} \right| + \frac{1}{2}|v|_{\text{Lip}} \left| \frac{x+1}{2} - \frac{y+1}{2} \right| \\
&= |v|_{\text{Lip}} \left| \frac{x}{2} - \frac{y}{2} \right| = \frac{1}{2}|v|_{\text{Lip}} |x-y|.
\end{aligned}$$

Therefore $|Pv|_{\text{Lip}} \leq \frac{1}{2}|v|_{\text{Lip}}$. Next note that for any $w : \Lambda \to \mathbb{R}$ Lipschitz, we have

$$\begin{aligned}
\left| w(x) - \int w \, d\mu \right| &\leq \int |w(x) - w(y)| \, d\mu(y) \leq |w|_{\text{Lip}} \int |x-y| \, d\mu(y) \\
&\leq |w|_{\text{Lip}} \sup_y |x-y| \, \mu(\Lambda) = |w|_{\text{Lip}} \sup_y |x-y|.
\end{aligned}$$

Hence

$$\begin{aligned}
\left\| P^n v - \int v \, d\mu \right\|_\infty &= \left\| P^n v - \int P^n v \, d\mu \right\|_\infty \leq \sup_x \Big(|P^n v|_{\text{Lip}} \sup_y |x-y| \Big) \\
&= |P^n v|_{\text{Lip}} \operatorname{diam} \Lambda = |P^n v|_{\text{Lip}} \leq |v|_{\text{Lip}} 2^{-n}.
\end{aligned}$$

Although mixing is exponential for this map with respect to Lipschitz observables, we can construct L^∞ observables that mix as slow as we desire:

Example 2.1.5. Let $\Sigma := \{0,1\}^{\mathbb{N}}$ be the space of sequences over two symbols and let $\sigma : \Sigma \to \Sigma$ be the left-shift, $\sigma\big((x_0 x_1 x_2 \cdots)\big) = (x_1 x_2 \cdots)$. Define the cylinders $[a_0 \cdots a_n; m] := \{x \in \Sigma : x_{m+i} = a_i, \, i = 0, \ldots, n\}$ and let μ be the $(1/2, 1/2)$-Bernoulli measure. That is, μ is the extension of the function that assigns value 2^{-n} to every cylinder of length n. Let $\gamma_n \geq 0$ be any sequence such that $\gamma_n \to 0$ and choose N such that $\gamma_N < 1/4$. Let $A = [1;0]$, $B = [0;N]$ and note

$$\left|\mu(\sigma^{-N} A \cap B) - \mu(A)\mu(B)\right| = \left|\mu([1;N] \cap [0;N]) - \frac{1}{4}\right| = \frac{1}{4} > \gamma_N.$$

2.2 Auxiliary Results

Here we collect some results that will prove useful for the rest of the thesis.

Proposition 2.2.1. *Let $(b_{r,n})_{r,n \geq 1}$ be a sequence of real numbers such that the following conditions hold:*

(a) $\sum_{r=1}^{\infty} \sup_n |b_{r,n}| < \infty$

(b) *for every $r \geq 1$,* $\lim_{n\to\infty} b_{r,n} = 0$.

Then

$$\sum_{r=1}^{\infty} b_{r,n} \to 0,$$

as $n \to \infty$.

Proof. Let $\epsilon > 0$. By condition (a) choose $R \geq 1$, independently of n, such that

$$\sum_{r=R+1}^{\infty} |b_{r,n}| < \epsilon.$$

For each $r \geq 1$, condition (b) guarantees us an N_r such that $|b_{r,n}| < \epsilon/R$ for all $n > N_r$. Therefore $\sum_{r=1}^{R} |b_{r,n}| < \epsilon$ for $n > \max_{r \leq R} N_r$. Hence

$$\sum_{r=1}^{\infty} |b_{r,n}| < 2\epsilon.$$

□

Proposition 2.2.2. *If c_n is a sequence of real numbers such that $n^{-1}c_n \to c$, then* $\sup_{0\leq t\leq 1} |n^{-1}c_{[nt]} - tc| \to 0$.

Proof. Suppose without loss that $c > 0$ and note that $-ntc \leq -[nt]c$. Calculate that

$$\sup_{0\leq t\leq 1} |n^{-1}c_{[nt]} - tc| = n^{-1} \sup_t |c_{[nt]} - ntc|$$
$$\leq n^{-1} \sup_t |c_{[nt]} - [nt]c| \leq n^{-1} \max_{0\leq j\leq n} |c_j - jc|.$$

Denoting $d_j = |c_j - jc|$, we see that $n^{-1}d_n \to 0$ and $\max_{0\le j\le n} d_j = d_{m(n)}$ for some non-decreasing sequence $m(n) \le n$. Note that $\sup_t |n^{-1}c_{[nt]} - tc| = n^{-1}d_{m(n)}$. Finally, either $m(n)$ is bounded and so $n^{-1}d_{m(n)} \to 0$ as $n \to \infty$, or

$$\sup_t |n^{-1}c_{[nt]} - tc| = n^{-1}d_{m(n)} \le m(n)^{-1}d_{m(n)} \to 0,$$

as $m(n) \to \infty$. □

Proposition 2.2.3. *Suppose $T : \Lambda \to \Lambda$ is an ergodic, measure-preserving map over the probability space $(\Lambda, \mathcal{A}, \mu)$, and let $f \in L^p(\Lambda)$ for some $p \ge 1$. Then*

$$n^{-1/p} \max_{0\le j\le n-1} |f| \circ T^j \to_{a.s.} 0.$$

Proof. By the ergodic theorem, we have

$$\frac{1}{n}(f^p)_n \to_{a.s.} \int f^p \, d\mu.$$

It follows that

$$\frac{1}{n} f^p \circ T^{n-1} = \frac{1}{n}\big((f^p)_n - (f^p)_{n-1}\big) = \frac{1}{n}(f^p)_n - \left(\frac{n}{n-1}\right)\frac{1}{n-1}(f^p)_{n-1} \to_{a.s.} 0.$$

Noting $\max_{0\le j\le n-1} |f|^p \circ T^j = \sup_{0\le t\le 1} |f|^p \circ T^{[(n-1)t]}$, Proposition 2.2.2 shows $n^{-1}\max_{0\le j\le n-1} |f|^p \circ T^j \to_{a.s.} 0$ and therefore $n^{-1/p}\max_{0\le j\le n-1} |f| \circ T^j \to_{a.s.} 0$. □

Remark 1. If T is invertible, then a near identical proof shows $n^{-1/p}\max_{1\le j\le n} |f| \circ T^{-j} \to_{a.s.} 0$.

We call a sequence of random variables $X_n : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ *tight* if for all $\epsilon > 0$ there exists a $K > 0$ such that $\max_{n\ge 1} \mathbb{P}\big(|X_n| \ge K\big) < \epsilon$.

Proposition 2.2.4. *Let $X_n, Y_n : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ be random variables such that X_n is a tight sequence and $Y_n \to_p 0$. Then $X_nY_n \to_p 0$.*

Proof. Let $\lambda, K > 0$ and note

$$\begin{aligned}\mathbb{P}(|X_nY_n| > \lambda) &= \mathbb{P}(|X_nY_n| > \lambda,\ |X_n| \le K) + \mathbb{P}(|X_nY_n| > \lambda,\ |X_n| > K)\\ &\le \mathbb{P}(|Y_n| \ge \frac{\lambda}{K}) + \mathbb{P}(|X_n| > K).\end{aligned}$$

Let $\epsilon > 0$ and choose K large enough so $\mathbb{P}(|X_n| > K) < \epsilon$. Now take n large enough so $\mathbb{P}(|Y_n| \geq \frac{\lambda}{K}) < \epsilon$. We now have

$$\mathbb{P}(|X_n Y_n| > \lambda) \leq 2\epsilon.$$

□

Proposition 2.2.5. *Define* $h : \mathbb{R}^n \to \mathbb{R}$, $h(a) := \max\{0, \sum_{i=1}^{1} a_i, \ldots, \sum_{i=1}^{n} a_i\}$. *Then*

$$|h(a) - h(b)| \leq \sum_{i=1}^{n} |a_i - b_i|.$$

Proof. If $h(a) = h(b) = 0$, then the result is clear. Next suppose $h(a) = \sum_{i=1}^{n_a} a_i$, $h(b) = \sum_{i=1}^{n_b} b_i$ for some $n_a, n_b \leq n$. Without loss assume $h(a) \geq h(b)$. Then

$$|h(a) - h(b)| = h(a) - h(b) = \sum_{i=1}^{n_a} a_i - \sum_{i=1}^{n_b} b_i.$$

By definition of h, we have $h(b) = \sum_{i=1}^{n_b} b_i \geq \sum_{i=1}^{p} b_i$ for any $p = 1, \ldots, n$. In particular, $-\sum_{i=1}^{n_b} b_i \leq -\sum_{i=1}^{n_a} b_i$ and so

$$|h(a) - h(b)| \leq \sum_{i=1}^{n_a} (a_i - b_i) \leq \sum_{i=1}^{n_a} |a_i - b_i| \leq \sum_{i=1}^{n} |a_i - b_i|.$$

□

Chapter 3

Limit Theorems for Non-Invertible Maps

Throughout this chapter, we suppose $T : \Lambda \to \Lambda$ is a ergodic, measure-preserving map over the probability space $(\Lambda, \mathcal{A}, \mu)$, with associated transfer operator P. Further assume $v \in L^\infty(\Lambda)$ is mean zero and satisfies the following mixing assumption:

$$\sum_{n=1}^{\infty} \|P^n v\|_1 < \infty. \tag{A}$$

Remark 1. If T is invertible, then Proposition 2.1.2 shows $P : L^p(\Lambda) \to L^p(\Lambda)$ is an isometry for every $p \geq 1$. Therefore Assumption (A) (and hence the hypothesis of Theorem 3.1.2) can never be non-trivially satisfied.

3.1 A Martingale Coboundary Decomposition

We say v admits a *martingale coboundary decomposition* if $v = m + \chi \circ T - \chi$ for some measurable observable χ with $m \in \operatorname{Ker} P$. We call $\chi \circ T - \chi$ the *coboundary* and say m and v are *cohomologous*. We call $\chi \circ T - \chi$ an $L^p(\Lambda)$ *coboundary* if $\chi \in L^p(\Lambda)$.

Proposition 3.1.1. *Let $k \geq 1$ and define $\chi_1^k := \sum_{n=1}^k P^n v$. We have the decomposition $v = m^{(k)} + \chi_1^k \circ T - \chi_1^k + P^k v$, where χ_1^k, $m^{(k)} \in L^\infty(\Lambda)$ and $m^{(k)} \in \operatorname{Ker} P$.*

Proof. For $k \geq 1$ define $m^{(k)} := v - \chi_1^k \circ T + \chi_1^k - P^k v$. Then $\chi_1^k, m^{(k)} \in L^\infty(\Lambda)$ follows easily since $\|P^n v\|_\infty \leq \|v\|_\infty < \infty$ for any $n \geq 1$. We see $m^{(k)} \in \operatorname{Ker} P$ by considering

$$\begin{aligned}Pm^{(k)} &= Pv - PU\chi_1^k + P\chi_1^k - P^{k+1}v = Pv - \chi_1^k + P\chi_1^k - P^{k+1}v \\ &= Pv - P^{k+1}v + \sum_{n=1}^{k}\big(P^{n+1}v - P^n v\big) = 0.\end{aligned}$$

□

By Assumption (A), $\chi := \sum_{n=1}^{\infty} P^n v$ is well defined. Further, $\chi_1^k \to \chi$ in $L^1(\Lambda)$ since $\|\chi - \chi_1^k\|_1 \le \sum_{n=k+1}^{\infty} \|P^n v\|_1 \to 0$.

Theorem 3.1.2. *We have the martingale coboundary decomposition $v = m + \chi \circ T - \chi$, where $\chi \in L^1(\Lambda)$. Further, $m \in \mathrm{Ker}(P)$, $m \in L^2(\Lambda)$ and $m^{(k)} \to m$ in $L^2(\Lambda)$.*

Proof. Define $m := v - \chi \circ T + \chi$. Then $\chi \in L^1(\Lambda)$ follows directly from Assumption (A) and we see $m \in \mathrm{Ker}\, P$ since

$$Pm = Pv - PU\chi + P\chi = Pv - \chi + P\chi = Pv + \sum_{n=1}^{\infty}\big(P^{n+1}v - P^n v\big) = 0.$$

It remains to show $m^{(k)} \to m$ in $L^2(\Lambda)$. Assumption (A) shows $m^{(k)} \to m$ in $L^1(\Lambda)$ since

$$\begin{aligned}\|m - m^{(k)}\|_1 &= \Big\|\chi_{k+1}^{\infty} - \chi_{k+1}^{\infty} \circ T + P^k v\Big\|_1 = \Big\|\sum_{n=k+1}^{\infty}\big(P^n v - UP^n v\big) + P^k v\Big\|_1 \\ &\le 2\sum_{n=k}^{\infty}\|P^n v\|_1 \to 0.\end{aligned}$$

By uniqueness of limits in $L^1(\Lambda)$ it suffices to show $m^{(k)}$ is a Cauchy sequence in $L^2(\Lambda)$. Without loss suppose $\ell > k$ and calculate that

$$\begin{aligned}m^{(\ell)} - m^{(k)} &= \chi_1^{\ell} - \chi_1^{\ell} \circ T - P^{\ell}v - \chi_1^k + \chi_1^k \circ T + P^k v \\ &= \chi_{k+1}^{\ell} - \chi_{k+1}^{\ell} \circ T + P^k v - P^{\ell} v = \chi_k^{\ell-1} - \chi_{k+1}^{\ell} \circ T.\end{aligned}$$

Hence, we see $m^{(k)}$ is a Cauchy sequence by considering

$$\int (m^{(\ell)} - m^{(k)})^2 \, d\mu = \int (m^{(\ell)} - m^{(k)})(\chi_k^{\ell-1} - \chi_{k+1}^{\ell} \circ T) \, d\mu$$
$$= \int \chi_k^{\ell-1} \big(m^{(\ell)} - m^{(k)}\big) \, d\mu - \int \chi_{k+1}^{\ell} \, P\big(m^{(\ell)} - m^{(k)}\big) \, d\mu = \int \chi_k^{\ell-1} \big(\chi_k^{\ell-1} - \chi_{k+1}^{\ell} \circ T\big) \, d\mu$$
$$= \int \big(\chi_k^{\ell-1}\big)^2 - \big(\chi_{k+1}^{\ell}\big)^2 \, d\mu = \int \left(\chi_k^{\ell-1} - \chi_{k+1}^{\ell}\right) \left(\chi_k^{\ell-1} + \chi_{k+1}^{\ell}\right) \, d\mu$$
$$= \int \left(P^k v - P^\ell v\right) \left(\chi_k^{\ell-1} + \chi_{k+1}^{\ell}\right) \, d\mu \leq \|P^k v - P^\ell v\|_\infty \|\chi_k^{\ell-1} + \chi_{k+1}^{\ell}\|_1$$
$$\leq 2\|v\|_\infty \big(\|\chi_k^{\ell-1}\|_1 + \|\chi_{k+1}^{\ell}\|_1\big) \ll \sum_{n=k}^{\ell} \|P^n v\|_1 \to 0,$$

as $\ell, k \to \infty$. $\square$

This technique is a standard part of the martingale method, first initialised by Gordin [18]. See also [21, 32].

Proposition 3.1.3 (L^2 orthogonality of martingale differences)**.** *Suppose $f, g : \Lambda \to \mathbb{R}$ are square integrable observables such that $f, g \in \operatorname{Ker} P$. Then for every $i, j \geq 0$,*

$$\int f \circ T^i \, g \circ T^j \, d\mu = \delta_{ij} \int f \, g \, d\mu,$$

where δ_{ij} is the Kronecker delta.

Proof. When $i = j$ the result follows from invariance of T. If $i < j$, then

$$\int f \circ T^i \, g \circ T^j \, d\mu = \int \big(f \, g \circ T^{j-i}\big) \circ T^i \, d\mu = \int f \, g \circ T^{j-i} \, d\mu$$
$$= \int Pf \, g \circ T^{j-i-1} \, d\mu = 0,$$

where we have used $f \in \operatorname{Ker} P$. The case $i > j$ follows from the same argument and the fact that $g \in \operatorname{Ker} P$. $\square$

For square integrable observables $f \in \operatorname{Ker} P$, Proposition 3.1.3 shows

$$\left\| \sum_{j=0}^{n-1} f \circ T^j \right\|_2 = n^{1/2} \|f\|_2.$$

3.2 Central Limit Theorem

3.2.1 Scalar-Valued Observables

We say a sequence of random variables $Z_n : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ satisfies a *central limit theorem* (CLT) with mean zero and variance σ^2 if

$$\mathbb{P}(Z_n \leq a) \to \int_{-\infty}^{a} (2\pi\sigma^2)^{-1/2} \exp\Big(-\frac{x^2}{2\sigma^2}\Big)\, dx,$$

for any $a \in \mathbb{R}$. This convergence is denoted by $Z_n \to_d \mathcal{N}(0,\sigma^2)$. The classical Lindeberg-Lévy CLT states that if $Z_n = n^{-1/2}\sum_{j=1}^{n} X_j$ for X_j a sequence of independent and identically distributed square integrable random variables with mean zero, then $Z_n \to_d \mathcal{N}(0,\sigma^2)$ with $\sigma^2 = \mathrm{Var}(X_1)$. We are interested in the case $X_j = v \circ T^j$, which form a dependent sequence.

Theorem 3.2.1 (CLT). *We have that $\sum_{r=1}^{\infty} \int v\ v \circ T^r\, d\mu$ converges and*

$$n^{-1/2} v_n \to_d \mathcal{N}(0,\sigma^2),$$

where $\sigma^2 = \int v^2\, d\mu + 2\sum_{r=1}^{\infty} \int v\ v \circ T^r\, d\mu$. Moreover, $\sigma^2 = 0$ if and only if v is an $L^1(\Lambda)$ coboundary.

Remark 1. This result is not new; a stronger result was shown by Liverani [32, Theorem 1.1].

Remark 2. Zweimüller [54, Corollary 1] proved that the CLT also holds with respect to every probability measure ν which is absolutely continuous with respect to μ.

We now give an example of a slowly mixing map which satisfies the hypothesis of Theorem 3.2.1.

Example 3.2.2. Suppose $0 < \alpha < 1$ and let $T : [0,1] \to [0,1]$ be an *intermittent map* of the form

$$Tx = \begin{cases} x(1 + 2^\alpha x^\alpha) & \text{if } 0 \leq x \leq 1/2 \\ 2x - 1 & \text{if } 1/2 < x \leq 1 \end{cases} \tag{3.1}$$

Figure 3.1 shows a graph of T for $\gamma = 0.9$. Liverani, Saussol and Vaienti [33, Lemma 2.3] showed there exists a unique ergodic invariant probability measure μ which is absolutely continuous with respect to Lebesgue. Let v be a mean zero Hölder observable. Young [53, Theorem 5(c)] showed $|C_n(v,w)| \leq Cn^{-1/\alpha+1}$ for all $w \in L^\infty(\Lambda)$. By Proposition 2.1.3, assumption (A) and therefore the CLT holds when $0 < \alpha < 1/2$.

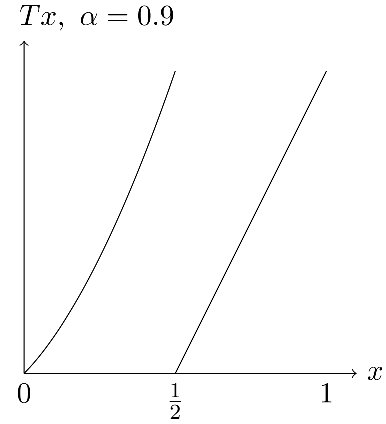

Figure 3.1: A graph of the intermittent map T defined by (3.1), for $\alpha = 0.9$.

When the correlations are not summable, the CLT generally fails [20, Theorem 3]: Take $1/2 < \alpha < 1$ in the above example (Example 3.2.2) and suppose v is Hölder with $v(0) \neq 0$. Then $n^{-\alpha} v_n \to_d Z$, where Z is a stable law [16, Chapter VI.1].

To prove Theorem 3.2.1, we first show that a CLT holds for the Birkhoff sums $n^{-1/2} m_n$ defined by the cohomologous observable m in Theorem 3.1.2, and then relate this to convergence of $n^{-1/2} v_n$. This technique was initialised by Gordin [18]. To show that the sequence $n^{-1/2} m_n$ satisfies a CLT, we show that it forms a martingale and apply a suitable martingale CLT result.

We call an increasing sequence of sub σ-algebras $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \subseteq \mathcal{A}$ a *filtration*. We say that a sequence of random variables Z_n is a *martingale* [15, Section 2.2] with respect to a filtration $\mathcal{A}_n$ if for each $n \geq 1$,

1. $Z_n \in L^1$
2. Z_n is $\mathcal{A}_n$-measurable
3. $\mathbb{E}(Z_{n+1} \,|\, \mathcal{A}_n) = Z_n$

We call X_j a sequence of *martingale differences* if $Z_n = \sum_{j=1}^{n} X_j$ is a martingale. In this case, condition 3 is equivalent to $\mathbb{E}(X_{j+1} \,|\, \mathcal{A}_j) = 0$ for all $j \geq 1$. The following proposition originates from Billingsley [5] (and proven independently by Ibragimov [25]).

Proposition 3.2.3 (Brown's Theorem [8]). *Let $f : X \to X$ be an ergodic, measure-preserving map, and let $\phi \in L^2(\Lambda)$ with $\mathbb{E}\phi = 0$. Suppose that $S_n := \sum_{j=0}^{n-1} \phi \circ f^j$ is a martingale. Then $n^{-1/2} S_n \to_d \mathcal{N}(0, \sigma^2)$, where $\sigma^2 = \mathbb{E}\phi^2$.* $\square$

For our Birkhoff sums $n^{-1/2}m_n$, the natural candidate for a filtration is $\mathcal{F}_n = T^{-n}\mathcal{A}$. However $\mathcal{F}_n$ forms a sequence of *decreasing* σ-algebras since $T^{-1}\mathcal{A} \subseteq \mathcal{A}$. The solution is to lift to an invertible map in the following way:

Proposition 3.2.4 ([44, pg. 13])**.** *Suppose that $T : (\Lambda, \mathcal{A}, \mu) \to (\Lambda, \mathcal{A}, \mu)$ is a surjective and measure-preserving map over the probability space $(\Lambda, \mathcal{A}, \mu)$. Then there exists a probability space $(\tilde{\Lambda}, \tilde{\mathcal{A}}, \tilde{\mu})$, an invertible, measure-preserving map $\tilde{T} : \tilde{\Lambda} \to \tilde{\Lambda}$ and a measure-preserving projection $\pi : \tilde{\Lambda} \to \Lambda$ such that $\pi \circ \tilde{T} = T \circ \pi$. If μ is ergodic, then $\tilde{\mu}$ can be taken to be ergodic too.* $\square$

We say the two systems $T, \tilde{T}$ in Proposition 3.2.4 are *semi-conjugate.* Using Proposition 3.2.4, define the lifted observable $\tilde{m} = m \circ \pi : \tilde{\Lambda} \to \mathbb{R}$, and the *forward and backward Birkhoff sums*

$$\tilde{m}_n := \sum_{j=0}^{n-1} \tilde{m} \circ \tilde{T}^j, \quad \tilde{m}_n^- := \sum_{j=1}^{n} \tilde{m} \circ \tilde{T}^{-j}. \tag{3.2}$$

Define the σ-algebras

$$\mathcal{A}_n := \tilde{T}^n \pi^{-1} \mathcal{A}.$$

Since $\mathcal{A}_{-1} = \tilde{T}^{-1}\mathcal{A}_0 = \tilde{T}^{-1}\tilde{\pi}^{-1}\mathcal{A} = \pi^{-1}T^{-1}\mathcal{A}$, we see $\mathcal{A}_{-1} \subseteq \mathcal{A}_0$. Hence $\mathcal{A}_n \subseteq \mathcal{A}_{n+1}$ for all $n \in \mathbb{Z}$ and $(\mathcal{A}_n)_{n\geq 0}$ defines a filtration.

Proposition 3.2.5. *Suppose $f \in L^1(\Lambda)$ is such that $f \in \operatorname{Ker} P$, and define $\tilde{f} := f \circ \pi$. Then $\tilde{f}_n^- := \sum_{j=1}^n \tilde{f} \circ \tilde{T}^{-j}$ is a martingale with respect to the filtration $\mathcal{A}_n = \tilde{T}^n \pi^{-1}\mathcal{A}$.*

Proof. Let $n \geq 1$. Firstly, we see $\|\tilde{f}_n^-\|_1 = n\|f\|_1 < \infty$. To see $\tilde{f}_n^- = \sum_{j=1}^n \tilde{f} \circ \tilde{T}^{-j}$ is $\mathcal{A}_n$-measurable, note that for every Borel subset $E \subseteq \mathbb{R}$ we have

$$\left(f \circ \pi \circ \tilde{T}^{-j}\right)^{-1} E = \tilde{T}^j \pi^{-1}(f^{-1}E) \in \mathcal{A}_j \subseteq \mathcal{A}_n.$$

It remains to show $\mathbb{E}\left(\tilde{f} \circ \tilde{T}^{-n} \,|\, \tilde{T}^{n-1}\pi^{-1}\mathcal{A}\right) = 0$. Note that

$$\begin{aligned}\mathbb{E}\left(\tilde{f} \circ \tilde{T}^{-n} \,|\, \tilde{T}^{n-1}\pi^{-1}\mathcal{A}\right) &= \mathbb{E}\left(\tilde{f} \,|\, \tilde{T}^{-1}\pi^{-1}\mathcal{A}\right) \circ \tilde{T}^{-n} = \mathbb{E}\left(f \circ \pi \,|\, \pi^{-1}T^{-1}\mathcal{A}\right) \circ \tilde{T}^{-n} \\ &= \mathbb{E}\left(f \,|\, T^{-1}\mathcal{A}\right) \circ \pi \circ \tilde{T}^{-n}.\end{aligned}$$

Proposition 2.1.2 shows $\mathbb{E}(f \,|\, T^{-1}\mathcal{A}) = (Pf) \circ T$ and therefore

$$\mathbb{E}\left(\tilde{f} \circ \tilde{T}^{-n} \,|\, \tilde{T}^{n-1}\pi^{-1}\mathcal{A}\right) = \mathbb{E}\left(f \,|\, T^{-1}\mathcal{A}\right) \circ \pi \circ \tilde{T}^{-n} = (Pf) \circ T \circ \pi \circ \tilde{T}^{-n} = 0,$$

as required. $\square$

Proposition 3.2.6 (Martingale CLT)**.** *The normalised Birkhoff sums $n^{-1/2}\tilde{m}_n^-$ satisfy a CLT with mean zero and variance $\sigma^2 = \int m^2\, d\mu$.*

Proof. Since m is a mean zero, $L^2(\Lambda)$ observable by Theorem 3.1.2 and π is measure-preserving, we see $\tilde{m}$ is a mean zero, $L^2(\tilde{\Lambda})$ observable. Proposition 3.2.5 shows that $\tilde{m}_n^-$ is a martingale with respect to the filtration $\mathcal{A}_n = \tilde{T}^n \pi^{-1} \mathcal{A}$. Proposition 3.2.3 now shows $n^{-1/2}\tilde{m}_n^- \to_d \mathcal{N}(0, \sigma^2)$, where $\sigma^2 = \int \tilde{m}^2\, d\tilde{\mu} = \int m^2\, d\mu$. □

In the following two propositions, we obtain the variance claimed in Theorem 3.2.1.

Proposition 3.2.7 (Green-Kubo Formula)**.** *The Green-Kubo formula holds:*

$$\lim_{n\to\infty} \frac{1}{n} \int v_n^2\, d\mu = \int v^2\, d\mu + 2\sum_{r=1}^{\infty} \int v\, v \circ T^r\, d\mu.$$

Proof. First note that

$$\frac{1}{n} \int v_n^2\, d\mu = \frac{1}{n} \sum_{i,j=0}^{n-1} \int v \circ T^i\, v \circ T^j\, d\mu.$$

We split the sum into the diagonal $i = j$, the lower triangle $i < j$ and upper triangle $i > j$. The lower and upper triangle are the same, so we obtain

$$\frac{1}{n} \int v_n^2\, d\mu = \int v^2\, d\mu + \frac{2}{n} \sum_{0\le i<j<n} \int v \circ T^i\, v \circ T^j\, d\mu.$$

Next calculate that

$$\begin{aligned}\frac{1}{n} \int v_n^2\, d\mu &= \int v^2\, d\mu + \frac{2}{n} \sum_{0\le i<j<n} \int \left(v\, v \circ T^{j-i}\right) \circ T^i\, d\mu \\ &= \int v^2\, d\mu + \frac{2}{n} \sum_{0\le i<j<n} \int v\, v \circ T^{j-i}\, d\mu \\ &= \int v^2\, d\mu + 2\sum_{r=1}^{\infty} \mathbb{1}(r \le n-1)(1 - r/n) \int v\, v \circ T^r\, d\mu. \qquad (3.3)\end{aligned}$$

Let $a_{r,n} := \mathbb{1}(r \le n-1)(1-r/n) \int v\, v \circ T^r\, d\mu$ and $a_r := \int v\, v \circ T^r\, d\mu$. We conclude that the limit $\lim_{n\to\infty} n^{-1} \int v_n^2\, d\mu$ exists and equals $\int v^2\, d\mu + 2\sum_{r=1}^{\infty} \int v\, v \circ T^r\, d\mu$ by showing $\sum_{r=1}^{\infty} a_{r,n} \to \sum_{r=1}^{\infty} a_r$, as $n \to \infty$. We verify the conditions of Proposition 2.2.1 with

$$b_{r,n} := a_{r,n} - a_r = (\mathbb{1}(r \le n-1)(1 - r/n) - 1) \int v\, v \circ T^r\, d\mu.$$

For condition (a), first note that $|b_{r,n}| \le 2\left|\int v\, v \circ T^r \, d\mu\right|$. The condition now follows from the following calculation:

$$\sum_{r=1}^{\infty} \sup_n |b_{r,n}| \le 2 \sum_{r=1}^{\infty} \left| \int P^r v\, v \, d\mu \right| \le 2\|v\|_\infty \sum_{r=1}^{\infty} \|P^r v\|_1 < \infty.$$

Condition (b) holds since for each fixed $r \ge 1$,

$$\lim_{n\to\infty} \mathbb{1}(r \le n-1)(1 - r/n) \int v\, v \circ T^r \, d\mu = \int v\, v \circ T^r \, d\mu.$$

Therefore

$$\lim_{n\to\infty} \frac{1}{n} \int v_n^2 \, d\mu = \int v^2 \, d\mu + 2\sum_{r=1}^{\infty} a_r = \int v^2 \, d\mu + 2 \sum_{r=1}^{\infty} \int v\, v \circ T^r \, d\mu.$$

□

Proposition 3.2.8. *We have that*

$$\int m^2 \, d\mu = \lim_{n\to\infty} \frac{1}{n} \int v_n^2 \, d\mu.$$

Proof. Let $m^{(k)} = v - \chi_1^k \circ T + \chi_1^k - P^k v$ be defined as in Proposition 3.1.1. Let $\epsilon > 0$ and choose k large enough so $\|m - m^{(k)}\|_2 < \epsilon$ and $\sum_{n=k}^{\infty} \|P^n v\|_1 < \epsilon^2$. By orthogonality of martingale differences (Proposition 3.1.3) we have

$$\begin{aligned}
\big|\|m\|_2 - n^{-1/2}\|v_n\|_2\big| &= \big|\|m\|_2 - \|m^{(k)}\|_2 + \|m^{(k)}\|_2 - n^{-1/2}\|v_n\|_2\big| \\
&= \big|\|m\|_2 - \|m^{(k)}\|_2 + n^{-1/2}\|m_n^{(k)}\|_2 - n^{-1/2}\|v_n\|_2\big| \\
&\le \|m - m^{(k)}\|_2 + n^{-1/2}\|m_n^{(k)} - v_n\|_2\big| \\
&\le \epsilon + n^{-1/2}\|v_n - m_n^{(k)}\|_2.
\end{aligned}$$

Next use the decomposition of v to calculate that

$$\begin{aligned}
\|v_n - m_n^{(k)}\|_2 &= \left\| \chi_1^k \circ T^n - \chi_1^k + \sum_{j=0}^{n-1} (P^k v) \circ T^j \right\|_2 \\
&\le 2\|\chi_1^k\|_2 + \left\| (P^k v)_n \right\|_2 \le 2k\|v\|_\infty + \left\| (P^k v)_n \right\|_2.
\end{aligned}$$

Since $P^k v \in L^\infty(\Lambda)$ and $\sum_{n=1}^{\infty} \|P^n(P^k v)\|_1 < \infty$, we use the Green-Kubo formula

(Proposition 3.2.7) with v replaced by $P^k v$ to obtain

$$\lim_{n\to\infty} \frac{1}{n} \int (P^k v)_n^2 \, d\mu = \int (P^k v)^2 \, d\mu + 2 \sum_{r=1}^{\infty} \int P^k v \, P^k v \circ T^r \, d\mu$$

$$\leq \|v\|_\infty \|P^k v\|_1 + 2 \sum_{r=1}^{\infty} \|v\|_\infty \|P^{r+k} v\|_1 \leq 2\|v\|_\infty \sum_{r=k}^{\infty} \|P^r v\|_1 < 2\|v\|_\infty \epsilon^2.$$

Therefore

$$\limsup_{n\to\infty} \left| \|m\|_2 - n^{-1/2} \|v_n\|_2 \right| \leq \limsup_{n\to\infty} \left(\epsilon + n^{-1/2}(2k\|v\|_\infty) + n^{-1/2} \|(P^k v)_n\|_2 \right)$$

$$\leq \left(1 + \sqrt{2\|v\|_\infty}\right)\epsilon,$$

and the result is shown. □

Proof of Theorem 3.2.1. By Proposition 3.2.6 we have

$$n^{-1/2} \tilde{m}_n^- \to_d \mathcal{N}(0, \sigma^2),$$

where $\sigma^2 = \int m^2 \, d\mu$. Propositions 3.2.7 and 3.2.8 show $\sigma^2 = \int v^2 \, d\mu + 2 \sum_{r=1}^{\infty} \int v \, v \circ T^r \, d\mu$. Since $\tilde{m}_n^- =_d \tilde{m}_n^- \circ \tilde{T}^n = \tilde{m}_n = m_n \circ \pi =_d m_n$, we have $n^{-1/2} m_n \to_d \mathcal{N}(0, \sigma^2)$. The central limit theorem for $n^{-1/2} v_n$ now follows easily, since

$$n^{-1/2} \|v_n - m_n\|_1 = n^{-1/2} \left\| \chi \circ T^n - \chi \right\|_1 \leq 2n^{-1/2} \|\chi\|_1 \to 0.$$

Finally, we see $\sigma^2 = \|m\|_2^2 = 0$ if and only if $m = 0$. □

3.2.2 Vector-Valued Observables

Suppose $v : \Lambda \to \mathbb{R}^d$ is a vector-valued observable, with each coordinate $v^i \in L^\infty(\Lambda)$. Define the $L^p(\Lambda, \mathbb{R}^d)$ norm

$$\|v\|_p := \big\| |v|_{\mathbb{R}^d} \big\|_{L^p(\Lambda, \mathbb{R})},$$

where $| \cdot |_{\mathbb{R}^d}$ is any norm on $\mathbb{R}^d$. We define the transfer operator P as acting coordinatewise on vector-valued observables. Assumption (A) becomes $\sum_{n=1}^{\infty} \|P^n v^i\|_1 < \infty$ for each $i = 1, \ldots, d$ and we immediately obtain the decompositions in Proposition 3.1.1 and Theorem 3.1.2 for vector-valued observables. Let v_n denote the n^{th} Birkhoff sum, and note $(v_n)^i = (v^i)_n$. For two vectors $f = (f^1, \ldots, f^d)$ and $g = (g^1, \ldots, g^d)$ define the outer product $(f \, g^T)_{ij} := f^i g^j$.

We say a sequence of random vectors $Z_n : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}^d$ satisfies a *central limit theorem* (CLT) with mean zero and covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$ if for any rectangle $I = [a_1, b_1] \times \cdots \times [a_d, b_d]$ we have

$$\mathbb{P}\big(Z_n \in I\big) \to \int_I \det(2\pi\Sigma)^{-1/2} \exp\big(-\frac{1}{2}x^T\Sigma^{-1}x\big)\; dx.$$

We denote this convergence by $Z_n \to_d \mathcal{N}(0, \Sigma)$. A result which will prove useful is the continuous mapping theorem.

Proposition 3.2.9 (Continuous Mapping Theorem [27, Lemma 4.3, Theorem 4.27])**.** *For any two metric spaces S and T, let $X_1, X_2, \ldots$ be random elements in S with $X_n \to_w X$, and consider a continuous map $f : S \to T$. Then $f(X_n) \to_w f(X)$. Moreover, if $X_n \to_p X$, then $f(X_n) \to_p f(X)$.* $\square$

A well known way of relating distributional convergence in $\mathbb{R}^d$ to distributional convergence in $\mathbb{R}$ is the Cramér-Wold device:

Proposition 3.2.10 (Cramér-Wold device [27, Corollary 4.5])**.** *Let X_n, X be random vectors in $\mathbb{R}^d$, for $n \geq 1$. Then $X_n \to_d X$ if and only if $c^T X_n \to_d c^T X$ in $\mathbb{R}$, for any $c \in \mathbb{R}^d$.*

For convergence in probability, the picture is a little simpler:

Proposition 3.2.11. *Let X_n, X be random vectors in $\mathbb{R}^d$, for $n \geq 1$. Then $X_n \to_p X$ if and only if $X_n^i \to_p X_n^i$ for each coordinate $i = 1, \ldots, d$.*

Proof. If $X_n \to_p X$, then $X_n^i \to_p X^i$ for each $i = 1, \ldots, d$ by the continuous mapping theorem. Next suppose $X_n^i \to_p X^i$ for every $i = 1, \ldots d$. Then we have

$$\begin{aligned}\mathbb{P}\left(\left|(X_n^1, \ldots, X_n^d) - (X^1, \ldots, X^d)\right| > a\right) &\leq \mathbb{P}\left(\sum_{i=1}^d \left|X_n^i - X^i\right| > a\right)\\ &\leq \sum_{i=1}^d \mathbb{P}\left(\left|X_n^i - X^i\right| > \frac{a}{d}\right) \to 0,\end{aligned}$$

as $a \to \infty$. $\square$

The main theorem for this section is the following:

Theorem 3.2.12 (Vector CLT [6, Theorem 29.5])**.** *We have that $\sum_{r=1}^\infty \int v\,\big(v \circ T^r\big)^T + (v \circ T^r)\, v^T\, d\mu$ converges and*

$$n^{-1/2} v_n \to_d \mathcal{N}(0, \Sigma),$$

where $\Sigma = \int v\, v^T\, d\mu + \sum_{r=1}^{\infty} \int v\, \left(v \circ T^r\right)^T + (v \circ T^r)\, v^T\, d\mu$*. For the degenerate case,* $\det \Sigma = 0$ *if and only if there exists a non-zero* $c \in \mathbb{R}^d$ *such that* $c^T v$ *is an* $L^1(\Lambda)$ *coboundary.*

The following two propositions give the correct covariance claimed in Theorem 3.2.12.

Proposition 3.2.13 (Vector Green-Kubo Formula)**.** *The following Green-Kubo formula holds:*

$$\lim_{n\to\infty} \frac{1}{n} \int v_n\, v_n^T\, d\mu = \int v\, v^T\, d\mu + \sum_{r=1}^{\infty} \int v\, \left(v \circ T^r\right)^T + \left(v \circ T^r\right) v^T\, d\mu.$$

Proof. We show the coordinate wise statement. That is, for each $i, j = 1, \dots, d$,

$$\lim_{n\to\infty} \frac{1}{n} \int v_n^i\, v_n^j\, d\mu = \int v^i v^j\, d\mu + \sum_{r=1}^{\infty} \int v^i\, v^j \circ T^r + v^j\, v^i \circ T^r\, d\mu.$$

Using the polarisation identity ($4 \int fg = \|f+g\|_2^2 - \|f-g\|_2^2$) we see

$$\frac{1}{n} \int v_n^i\, v_n^j\, d\mu = \frac{1}{4n} \int (v_n^i + v_n^j)^2\, d\mu - \frac{1}{4n} \int (v_n^i - v_n^j)^2\, d\mu. \tag{3.4}$$

Since $v^i \pm v^j \in L^\infty(\Lambda)$ and $\sum_{k=1}^{\infty} \|P^k (v^i \pm v^j)\|_1 \le \sum_{k=1}^{\infty} \|P^k v^i\|_1 + \|P^k v^j\|_1 < \infty$ we use the scalar Green-Kubo formula (Proposition 3.2.7) with v replaced by $v^i \pm v^j$ to see

$$\lim_{n\to\infty} \frac{1}{4n} \int (v^i \pm v^j)_n\, d\mu = \frac{1}{4} \int (v^i \pm v^j)^2\, d\mu + \frac{2}{4} \sum_{r=1}^{\infty} \int (v^i \pm v^j)\ (v^i \pm v^j) \circ T^r\, d\mu.$$

Therefore, using the polarisation identity once more and taking $\lim_{n\to\infty}$ in (3.4) gives

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} \int v_n^i\, v_n^j\, d\mu \\
&= \int v^i\, v^j\, d\mu + \frac{2}{4} \sum_{r=1}^{\infty} \int (v^i + v^j)\, (v^i + v^j) \circ T^r - (v^i - v^j)\, (v^i - v^j) \circ T^r\, d\mu \\
&= \int v^i\, v^j\, d\mu + \sum_{r=1}^{\infty} \int v^i\, v^j \circ T^r + v^j\, v^i \circ T^r\, d\mu.
\end{aligned}$$

$\square$

Proposition 3.2.14. *We have that*

$$\int m\, m^T \, d\mu = \lim_{n\to\infty} \frac{1}{n} \int v_n\, v_n^T \, d\mu.$$

Proof. We show the coordinate wise statement. That is, for each $i, j = 1, \ldots, d$,

$$\lim_{n\to\infty} \frac{1}{n} \int v_n^i\, v_n^j \, d\mu = \int m^i\, m^j \, d\mu.$$

Using the polarisation identity ($4 \int fg = \|f+g\|_2^2 - \|f-g\|_2^2$) we see

$$\frac{4}{n} \int v_n^i\, v_n^j \, d\mu = \frac{1}{n} \int (v_n^i + v_n^j)^2 - (v_n^i - v_n^j)^2 \, d\mu = \frac{1}{n} \int (v^i + v^j)_n^2 - (v^i - v^j)_n^2 \, d\mu.$$

Note we have the martingale coboundary decompositions

$$v^i \pm v^j = \left(m^i \pm m^j\right) + \left(\chi^i \pm \chi^j\right) \circ T - \left(\chi^i \pm \chi^j\right),$$

for the scalar-valued observables $v^i \pm v^j$. Since $v^i \pm v^j \in L^\infty(\Lambda)$ and $\sum_{k=1}^\infty \|P^k(v^i \pm v^j)\|_1 \le \sum_{k=1}^\infty \|P^k v^i\|_1 + \|P^k v^j\|_1 < \infty$ we use the scalar case (Proposition 3.2.8) with v replaced by $v^i \pm v^j$ to see

$$\begin{aligned} &\lim_{n\to\infty} \frac{4}{n} \int v_n^i\, v_n^j \, d\mu = \lim_{n\to\infty} \frac{1}{n} \int (v^i + v^j)_n^2 - \lim_{n\to\infty} \frac{1}{n} \int (v^i - v^j)_n^2 \, d\mu \\ &= \int (m^i + m^j)^2 - (m^i - m^j)^2 \, d\mu = 4 \int m^i\, m^j \, d\mu. \end{aligned}$$

□

Proof of Theorem 3.2.12. By the Cramér-Wold device (Theorem 3.2.10) it suffices to show $c^T(n^{-1/2} v_n) = n^{-1/2}(c^T v)_n \to_d c^T \mathcal{N}(0, \Sigma) = \mathcal{N}(0, c^T \Sigma c)$ for any $c \in \mathbb{R}^d$. Note that $c^T v \in L^\infty(\Lambda)$, $\int c^T v \, d\mu = 0$ and $\sum_{n=1}^\infty \|P^n(c^T v)\|_1 < \infty$. By the scalar CLT (Theorem 3.2.1), $n^{-1/2}(c^T v)_n \to_d \mathcal{N}(0, \sigma^2)$ with $\sigma^2 = \lim_{n\to\infty} n^{-1} \int (c^T v)_n^2 \, d\mu$. Next, we see $\sigma^2 = c^T \Sigma c$, since

$$\begin{aligned} c^T \Sigma c &= c^T \left(\lim_{n\to\infty} \frac{1}{n} \int v_n\, v_n^T \, d\mu \right) c = \lim_n \frac{1}{n} \int c^T (v_n\, v_n^T)\; c \, d\mu \\ &= \lim_n \frac{1}{n} \int (c^T\, v_n)(v_n^T\, c) \, d\mu = \lim_n \frac{1}{n} \int (c^T\, v)_n^2 \, d\mu = \sigma^2. \end{aligned}$$

Proposition 3.2.13 now shows $\Sigma = \int v\, v^T \, d\mu + \sum_{r=1}^\infty \int v \left(v \circ T^r\right)^T + \left(v \circ T^r\right) v^T \, d\mu$.

For the degenerate case, first suppose $\det \Sigma = 0$. Then there exists a non-zero $c \in \mathbb{R}^d$ such that $\sigma^2 = \Sigma c = 0$ and hence $c^T \Sigma c = 0$. From the above calculation we have $n^{-1/2}(c^T v)_n \to_d \mathcal{N}(0, c^T \Sigma c)$ and so by the scalar CLT (Theorem 3.2.1), $c^T v$ is an $L^1(\Lambda)$ coboundary. Finally, suppose there exists a non-zero $c \in \mathbb{R}^d$ such that $c^T v$ is an $L^1(\Lambda)$ coboundary. Then $c^T \Sigma c = 0$ by the scalar CLT. Since Σ is positive semi-definite we have $\Sigma = A^2$ for some symmetric matrix A. Then $0 = c^T \Sigma c = c^T A^T A c = \|Ac\|^2$ and so $Ac = 0$. Therefore $\det A = 0$ and $\det \Sigma = (\det A)^2 = 0$. $\square$

3.3 Weak Invariance Principle

Here we consider vector-valued observables $v : \Lambda \to \mathbb{R}^d$. Define the process

$$W_n : [0,1] \to \mathbb{R}^d, \quad W_n(t) := n^{-1/2} v_{[nt]} = n^{-1/2} \sum_{j=0}^{[nt]-1} v \circ T^j, \tag{3.5}$$

with the convention that $\sum_{j=0}^{-1} v \circ T^j = 0$. Figure 3.2 illustrates a typical sample path of W_n for $d = 1$.

Suppose $X_n : (\Lambda, \mathcal{A}, \mu) \to S$, $n \geq 1$, and $X : (\Omega, \mathcal{F}, \mathbb{P}) \to S$ are random elements of some metric space S. We say X_n *converges weakly* to X, denoted $X_n \to_w X$, if $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$ for every bounded, continuous function $f : S \to \mathbb{R}$ [7, Section 3].

One choice for S is the *Skorohod space* $D([0,1], \mathbb{R}^d)$, the space of $\mathbb{R}^d$-valued functions with left limits that are right continuous. We call a function $X \in D([0,1], \mathbb{R}^d)$ a *càdlàg process* or *càdlàg function*. Any càdlàg process is bounded [7, Lemma 12.1], and we can endow the Skorohod space $D([0,1], \mathbb{R}^d)$ with the supremum norm $\|\cdot\|_\infty$ to obtain a non-separable Banach space [7, Section 15]. Clearly $C([0,1], \mathbb{R}^d) \subset D([0,1], \mathbb{R}^d)$. We have $W_n \in D([0,1], \mathbb{R}^d)$. An important fact that we have already used is the following:

Proposition 3.3.1 ([7, Theorem 3.1])**.** *Suppose $X_n, Y_n : S \to \mathbb{R}$, $n \geq 1$, and $X, Y : S \to \mathbb{R}$ are random elements of a separable metric space S. Assume that $X_n \to_w X$, $Y_n \to_p Y$ and that Y is constant almost surely. Then $X_n + Y_n \to_w X + Y$.*

Since $\big(D([0,1], \mathbb{R}^d), \|\cdot\|_\infty\big)$ is not separable, we require a workaround to use this proposition: We can endow $D([0,1], \mathbb{R}^d)$ with a weaker topology, the *Skorohod topology* $\mathcal{J}_1$. The Skorohod space $(D([0,1], \mathbb{R}^d), \mathcal{J}_1)$ is separable and complete under this topology [7, Theorem 12.2].

Proposition 3.3.2 ([7, Section 15])**.** *If $X_n \to_w X$ in $(D([0,1], \mathbb{R}^d), \|\cdot\|_\infty)$, then convergence also holds in $(D([0,1], \mathbb{R}^d), \mathcal{J}_1)$. If $X_n \to_w X$ in $(D([0,1], \mathbb{R}^d), \mathcal{J}_1)$ and $X \in C([0,1], \mathbb{R}^d)$, then $X_n \to_w X$ in $(D([0,1], \mathbb{R}^d), \|\cdot\|_\infty)$.*

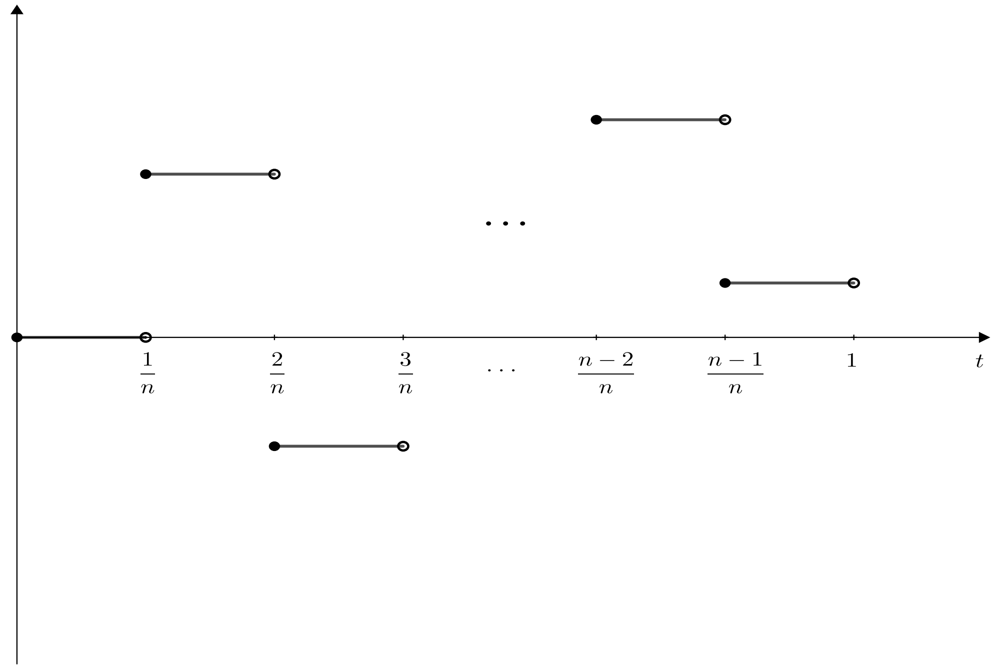

Figure 3.2: A typical sample path of the process W_n for $d = 1$.

Remark. This proposition allows us to work in the $\mathcal{J}_1$ topology when required, and 'upgrade' convergence to the $\|\cdot\|_\infty$ topology.

We call $W : (\Omega, \mathcal{F}, \mathbb{P}) \to C([0,1], \mathbb{R}^d)$ a *d dimensional Brownian motion* with covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$ if $W(0) = 0$ and

1. *(Independent increments)* The increments $W(t_1) - W(t_0), \ldots, W(t_k) - W(t_{k-1})$ are independent for all $0 = t_0 \le t_1 \le \cdots \le t_k \le 1$.

2. *(Normal increments)* For all $0 \le s \le t \le 1$ we have $W(t) - W(s) =_d \mathcal{N}(0, \Sigma(t-s))$.

We say a sequence of random functions Z_n satisfies a *weak invariance principle* (WIP) if $Z_n \to_w W$ in $\big(C([0,1], \mathbb{R}^d), \|\cdot\|_\infty\big)$ or $\big(D\big([0,1], \mathbb{R}^d\big), \|\cdot\|_\infty\big)$.

Theorem 3.3.3 (WIP)**.** *We have the convergence*

$$W_n \to_w W$$

in $\big(D\big([0,1], \mathbb{R}^d\big), \|\cdot\|_\infty\big)$*, where* W *is* d *dimensional Brownian motion with mean zero and covariance* $\Sigma = \int v\, v^T\, d\mu + \sum_{r=1}^{\infty} \int v\, \big(v \circ T^r\big)^T + \big(v \circ T^r\big)\, v^T\, d\mu$.

Remark 1. This was first proven by Dedecker and Rio [11, Theorem 1] under similar hypotheses: They work with invertible maps $T : (\Lambda, \mathcal{A}, \mu) \to (\Lambda, \mathcal{A}, \mu)$ and assume existence of a sub-sigma algebra $\mathcal{A}_0 \subset \mathcal{A}$ such that $T^{-1}\mathcal{A}_0 \supset \mathcal{A}_0$. Their observable $v : \Lambda \to \mathbb{R}$ is taken to be square integrable and $\mathcal{A}_0$-measurable. The measurability assumption is an important one, and is what causes their result to be analogous to the non-invertible case. They also assume the mixing condition

$$\sum_{n=0}^{\infty} v\,\mathbb{E}(v \circ T^{-n} \,|\, \mathcal{A}_0) \quad \text{converges in } L^1(\Lambda).$$

For more details, see the discussion at the beginning of Chapter 4 (note that the first term in Assumption ($\mathrm{A}_{\mathrm{inv}}$) dissapears when v is $\mathcal{A}_0$-measurable. Also, Dedecker & Rio have time running backwards with respect to our convention).

Remark 2. As in Theorem 3.2.12, $\det \Sigma = 0$ if and only if there exists a non-zero $c \in \mathbb{R}^d$ such that $c^T v$ is an $L^1(\Lambda)$ coboundary.

Remark 3. We could equally define our processes in $D([0,K], \mathbb{R}^d)$ for any integer $K \geq 1$, and the result still holds. Hence the result is true if we define our processes in $D([0,\infty), \mathbb{R}^d)$ by [7, Section 16].

Remark 4. Zweimüller [54, Corollary 3] proved that the WIP also holds with respect to every probability measure ν which is absolutely continuous with respect to μ.

Example 3.3.4. For $0 < \alpha < 1$, let $T : [0,1] \to [0,1]$ be the intermittent map defined by equation (3.1), introduced in Example 3.2.2. Let $v : [0,1] \to \mathbb{R}$ be a mean zero Hölder observable. Recall $\|P^n v\|_1 \leq C n^{-1/\alpha+1}$ and so Theorem 3.3.3 gives weak convergence of the process W_n to Brownian motion W in $\big(D\big([0,1], \mathbb{R}\big), \|\cdot\|_\infty\big)$ for $0 < \alpha < 1/2$.

For the above example (Example 3.3.4), the WIP can fail if the correlations are not summable: If $1/2 < \alpha < 1$ and $v : \Lambda \to \mathbb{R}$ is a mean-zero observable such that $v(0) \neq 0$, then Melbourne and Zweimüller [41, Theorem 1.1] proved that the process $n^{-\alpha} v_{[nt]}$ converges to an α-stable Lévy process [27, Pg. 291–292] in some suitable topology over $D\big([0,1], \mathbb{R}\big)$.

To prove Theorem 3.3.3, we follow a similar procedure as in Section 3.2. First, we setup and state the martingale WIP (Proposition 3.3.5). Next, we show that the lifted backwards process satisfies the martingale WIP. Finally, we prove that the lifted backwards process has the same weak limit as W_n.

We call a sequence of sigma algebras $\{\mathcal{A}_t \,|\, 0 \leq t \leq 1\}$ a *filtration* if for every $0 \leq s \leq t \leq 1$ we have $\mathcal{A}_s \subseteq \mathcal{A}_t$. A continuous time stochastic process $\{X(t) : \Omega \to \mathbb{R}^d \,|\, 0 \leq t \leq 1\}$ is said to be a *adapted* to a filtration $\mathcal{F}_t$ if X_t is $\mathcal{F}_t$-measurable for every $t \in [0,1]$. We say X_t is a *martingale* [15, Section 2.2] with respect to the filtration $\{\mathcal{A}_t \,|\, 0 \leq t \leq 1\}$ if for all $0 \leq s \leq t \leq 1$ we have

1. $X(t)$ is $\mathcal{A}_t$-measurable

2. $\|X(t)\|_1 < \infty$

3. $\mathbb{E}(X(t) \,|\, \mathcal{A}_s) = X(s)$ a.s.

For $X \in D([0,1], \mathbb{R})$, define $X(t-) := \lim_{s \nearrow t} X(s)$. Let

$$J(X,t) := \sup \big\{ |X(s) - X(s-)| : 0 < s \le t \big\}$$

be the *largest jump* up to time t. Let $t \in [0,1]$ and suppose Δ_ℓ is a sequence of refining partitions of $[0,t]$ such that $\max_{\Delta_\ell} |t_{k+1} - t_k| \to 0$ as $\ell \to \infty$. For $Y \in D([0,1], \mathbb{R})$, define the *quadratic covariation* between X and Y as

$$[X,Y](t) := \lim_{\ell \to \infty} \sum_{k=0}^{\ell-1} \big(X(t_{k+1}) - X(t_k)\big)\big(Y(t_{k+1}) - Y(t_k)\big),$$

where the limit is taken in probability, whenever it exists. Let $\Delta X(t) := X(t) - X(t-)$. If X and Y are piecewise constant, then

$$[X,Y](t) = X(0)Y(0) + \sum_{0 \le s \le t} \Delta X(s) \Delta Y(s). \tag{3.6}$$

The statement of the following proposition has been taken from [51, Theorem 2.1], which is an adaptation of [15, Section 7.1].

Proposition 3.3.5 (Martingale WIP [51, Theorem 2.1])**.** *For each $n \ge 1$, let $F_n \in D([0,1], \mathbb{R}^d)$, and suppose $F_n := (F_n^1, \dots, F_n^d)$ is a martingale with respect to the filtration $\{\mathcal{A}_{n,t} \,|\, 0 \le t \le 1\}$ such that $F_n(0) = 0$. Let Σ be a $d \times d$ symmetric and positive semi-definite matrix of real numbers. Assume*

(i) for each $0 \le t \le 1$, we have $\lim_{n\to\infty} \mathbb{E}(J(F_n, t)) = 0$

(ii) for each coordinate $\beta, \gamma = 1, \dots, d$ and every $0 \le t \le 1$ we have $[F_n^\beta, F_n^\gamma](t) \to_d t\Sigma_{\beta\gamma}$

Then

$$F_n \to_w W$$

in $\big(D([0,1], \mathbb{R}^d), \|\cdot\|_\infty\big)$, where W is a d dimensional Brownian motion with mean zero and covariance Σ. $\square$

By Proposition 3.2.4, lift the probability space $(\Lambda, \mathcal{A}, \mu)$ and map T to obtain a probability space $(\tilde{\Lambda}, \tilde{\mathcal{A}}, \tilde{\mu})$, where $\tilde{\mu}$ is an ergodic, invariant measure with respect to the invertible map $\tilde{T} : \tilde{\Lambda} \to \tilde{\Lambda}$. Associated to this is a measure-preserving projection $\pi : \tilde{\Lambda} \to \Lambda$ such that $\pi \circ \tilde{T} = T \circ \pi$. Define the lifted observable $\tilde{v} := v \circ \pi : \tilde{\Lambda} \to \mathbb{R}^d$ and similarly define $\tilde{m} := m \circ \pi$, $\tilde{m}^{(k)} := m^{(k)} \circ \pi$. Recall that $\|\tilde{v}\|_p = \|v\|_p$ for every $p \geq 1$. Define the forwards lifted process

$$\tilde{W}_n : [0,1] \to \mathbb{R}^d, \quad \tilde{W}_n(t) := n^{-1/2} \sum_{j=0}^{[nt]-1} \tilde{v} \circ \tilde{T}^j,$$

and the backwards lifted process

$$\tilde{W}_n^- : [0,1] \to \mathbb{R}^d, \quad \tilde{W}_n^-(t) := n^{-1/2} \sum_{j=1}^{[nt]} \tilde{v} \circ \tilde{T}^{-j}.$$

Define the backwards lifted processes $\tilde{M}_n^-, \tilde{M}_n^{(k),-}$ in a similar fashion with $\tilde{m}, \tilde{m}^{(k)}$ respectively, and $\sum_1^0 \tilde{v} \circ \tilde{T}^{-j} = \sum_1^0 \tilde{m} \circ \tilde{T}^{-j} = \sum_1^0 \tilde{m}^{(k)} \circ \tilde{T}^{-j} = 0$.

Proposition 3.3.6. *The quadratic variation of $\tilde{M}_n^-$ is given by*

$$[\tilde{M}_n^{\beta,-}, \tilde{M}_n^{\gamma,-}](t) = \frac{1}{n} \sum_{j=1}^{[nt]} (\tilde{m}^\beta \tilde{m}^\gamma) \circ \tilde{T}^{-j},$$

for each $\beta, \gamma = 1, \ldots, d$.

Proof. Note that the jumps of $\tilde{M}_n^{\beta,-}(t)$ occur at the points $t = \frac{j}{n}$, for $j = 1, \ldots, [nt]$. Moreover, each jump is given by $\Delta \tilde{M}_n^{\beta,-}(\frac{j}{n}) = n^{-1/2} \tilde{m}^\beta \circ \tilde{T}^{-j}$. The result now follows from (3.6). $\square$

Proposition 3.3.7. *Suppose $f \in L^1(\Lambda, \mathbb{R}^d)$ is such that $f \in \operatorname{Ker} P$, and define $\tilde{f} := f \circ \pi$. Then $\tilde{f}_n^-(t) := \sum_{j=1}^{[nt]} \tilde{f} \circ \tilde{T}^{-j}$ is a martingale with respect to the filtration $\mathcal{A}_{n,t} = \{\tilde{T}^{[nt]} \pi^{-1} \mathcal{A} \,|\, 0 \leq t \leq 1\}$.*

Proof. Let $n \geq 1$. Firstly, $\tilde{f}_n^-(t) \in L^1(\tilde{\Lambda}, \mathbb{R}^d)$ since

$$\|\tilde{f}_n^-(t)\|_1 = \left\| \sum_{j=1}^{[nt]} \tilde{f} \circ \tilde{T}^{-j} \right\|_1 \leq [nt] \|f\|_1 < \infty.$$

To see $\tilde{f}_n^-(t) = \sum_{j=1}^{[nt]} \tilde{f} \circ \tilde{T}^{-j}$ is $\mathcal{A}_{n,t}$-measurable, note that for any Borel subset E of $\mathbb{R}^d$ and all $s \leq t$ we have

$$\left(f \circ \pi \circ \tilde{T}^{-[ns]}\right)^{-1} E = \tilde{T}^{[ns]} \pi^{-1}(f^{-1} E) \in \mathcal{A}_{n,s} \subseteq \mathcal{A}_{n,t}.$$

It remains to show

$$\mathbb{E}(\tilde{f}_n^-(t) \mid \mathcal{A}_{n,s}) = \tilde{f}_n^-(s),$$

for all $s \le t$. If $[nt] = [ns]$, then $\tilde{f}_n^-(t) = \tilde{f}_n^-(s)$ and we are done. Suppose $[nt] > [ns]$ and note that

$$\mathbb{E}\Bigg(\sum_{j=1}^{[nt]} \tilde{f} \circ \tilde{T}^{-j} \mid \mathcal{A}_{n,s}\Bigg) = \tilde{f}_n^-(s) + \sum_{j=[ns]+1}^{[nt]} \mathbb{E}(\tilde{f} \circ \tilde{T}^{-j} \mid \tilde{T}^{[ns]}\pi^{-1}\mathcal{A}).$$

Let $j = [ns]+1, \ldots, [nt]$ be fixed, and note it suffices to show $\mathbb{E}(\tilde{f} \circ \tilde{T}^{-j} \mid \tilde{T}^{[ns]}\pi^{-1}\mathcal{A}) = 0$. Write $\tilde{T}^{[ns]}\pi^{-1}\mathcal{A} = \tilde{T}^{j-k}\pi^{-1}\mathcal{A}$, where $k = j - [ns] > 0$. Since $f \in \operatorname{Ker} P$,

$$\begin{aligned}
&\mathbb{E}(\tilde{f} \circ \tilde{T}^{-j} \mid \tilde{T}^{j-k}\pi^{-1}\mathcal{A}) = \mathbb{E}(\tilde{f} \mid \tilde{T}^{-k}\pi^{-1}\mathcal{A}) \circ \tilde{T}^{-j} = \mathbb{E}(f \circ \pi \mid \pi^{-1}T^{-k}\mathcal{A}) \circ \tilde{T}^{-j} \\
&= \mathbb{E}(f \mid T^{-k}\mathcal{A}) \circ \pi \circ \tilde{T}^{-j} = (P^k f) \circ T^k \circ \pi \circ \tilde{T}^{-j} = 0.
\end{aligned}$$

□

Proposition 3.3.8 (Martingale WIP)**.** *We have*

$$\tilde{M}_n^- \to_w W$$

in $\big(D([0,1], \mathbb{R}^d), \|\cdot\|_\infty\big)$*, where* W *is a* d *dimensional Brownian motion with mean zero and covariance* $\Sigma = \int m\, m^T \, d\mu$*.*

Proof. By Proposition 3.3.7, $\tilde{M}_n^-$ is a martingale. Since $\tilde{M}_n^-(0) = 0$, it remains to verify conditions (i) and (ii) of Proposition 3.3.5. We begin by noting that for $0 \le t \le 1$ we have

$$J(\tilde{M}_n^-, t) = n^{-1/2} \max_{1 \le j \le [nt]} |\tilde{m} \circ \tilde{T}^{-j}| \le n^{-1/2} \max_{1 \le j \le n} |\tilde{m} \circ \tilde{T}^{-j}|$$

For condition (i), we consider the square of the expected value of the jumps:

$$\Big[\mathbb{E}(J(\tilde{M}_n^-, t))\Big]^2 = \|J(\tilde{M}_n^-, t)\|_1^2 \le \|J(\tilde{M}_n^-, t)\|_2^2 \le \int \max_{1 \le j \le n} n^{-1}\tilde{m}^2 \circ \tilde{T}^{-j} \, d\tilde{\mu}. \quad (3.7)$$

We use the fact that for any $a > 0$ we have

$$n^{-1}\tilde{m}^2 \circ \tilde{T}^{-j} \le a^2 + n^{-1}\tilde{m}^2 \circ \tilde{T}^{-j}\, \mathbb{1}\{n^{-1}\tilde{m}^2 \circ \tilde{T}^{-j} \ge a^2\}. \quad (3.8)$$

Take the maximum over $j = 1, \dots, n$ and integrate in (3.8) to obtain

$$\int \max_{1 \le j \le n} n^{-1} \tilde{m}^2 \circ \tilde{T}^{-j} \, d\tilde{\mu} \le a^2 + n^{-1} \int \max_{1 \le j \le n} \tilde{m}^2 \circ \tilde{T}^{-j} \, \mathbb{1}\{\tilde{m}^2 \circ \tilde{T}^{-j} \ge na^2\} \, d\tilde{\mu}$$
$$\le a^2 + n^{-1} \int \sum_{j=1}^{n} \tilde{m}^2 \circ \tilde{T}^{-j} \mathbb{1}\{\tilde{m}^2 \circ \tilde{T}^{-j} \ge na^2\} \, d\tilde{\mu}$$
$$= a^2 + \int \tilde{m}^2 \, \mathbb{1}\{\tilde{m}^2 \ge na^2\} \, d\tilde{\mu}. \tag{3.9}$$

Combining (3.7) and (3.9), and taking $\limsup_{n\to\infty}$ shows

$$\limsup_{n\to\infty} \Big[\mathbb{E}(J(\tilde{M}_n^-, t))\Big]^2 \le a^2.$$

Since $a > 0$ was arbitrary we see condition (i) is satisfied.

Condition (ii) follows from Proposition 3.3.6 and the ergodic theorem, since

$$[\tilde{M}_n^{\beta,-}, \tilde{M}_n^{\gamma,-}](t) = \frac{1}{n} \sum_{j=1}^{[nt]} (\tilde{m}^\beta \tilde{m}^\gamma) \circ \tilde{T}^{-j}$$
$$= \frac{[nt]}{n} \frac{1}{[nt]} \sum_{j=1}^{[nt]} (\tilde{m}^\beta \tilde{m}^\gamma) \circ \tilde{T}^{-j} \to_{a.s.} t \int m^\beta \, m^\gamma \, d\mu = t\Sigma_{\beta\gamma}.$$

Therefore Proposition 3.3.5 gives the convergence $\tilde{M}_n^- \to_w W$ in $\big(D([0,1], \mathbb{R}^d), \|\cdot\|_\infty\big)$, where $\Sigma = \int \tilde{m}\, \tilde{m}^T \, d\mu = \int m \, m^T \, d\mu$. ☐

For a sequence $S_n = \sum_{j=1}^n X_j$ of random vectors X_j define $\overline{S_n} = \max_{1 \le \ell \le n} |S_\ell|$. Note that

$$\overline{v_n} = \max_{1 \le \ell \le n} |v_\ell| = \max_{1 \le \ell \le n} \left| \sum_{j=1}^{\ell} v \circ T^j \right|.$$

To show $\tilde{M}_n^-$ has the same weak limit as W_n, we will need the two following maximal inequalities.

Proposition 3.3.9 (Doob's inequality [21, Theorem 2.2])**.** *Let S_n, $n \ge 1$ be a sequence of square integrable random variables that form a martingale. Then*

$$\|S_n\|_2 \le \|\overline{S_n}\|_2 \le 2\|S_n\|_2.$$

☐

Proposition 3.3.10 (Rio's inequality [48])**.** *Let X_j be a sequence of square integrable random variables adapted to some filtration $\mathcal{F}_j$, for $j \ge 1$, and define*

$S_n := \sum_{j=1}^n X_j$. *Then*

$$\mathbb{E}\big(\overline{S_n}^2\big) \le 16 \sum_{j=1}^{n} \max_{j\le \ell\le n} b_{j,\ell},$$

where $b_{j,\ell} = \Big\| X_j \sum_{u=j}^{\ell} \mathbb{E}\big(X_u | \mathcal{F}_j\big)\Big\|_1$. □

Proposition 3.3.11. *We have that $\tilde{W}_n^-$ has the same weak limit as $\tilde{M}_n^-$, so $\tilde{W}_n^- \to_w W$ in $\big(D([0,1],\mathbb{R}^d), \|\cdot\|_\infty\big)$, where W is a d dimensional Brownian motion with mean zero and covariance $\Sigma = \int m\, m^T\, d\mu$.*

Proof. Proposition 3.3.8 shows $\tilde{M}_n^- \to_w W$, where W is a d dimensional Brownian motion with mean zero and covariance $\Sigma = \int m\, m^T\, d\mu$. Therefore, it suffices to show

$$\sup_{0\le t\le 1} |\tilde{M}_n^-(t) - \tilde{W}_n^-(t)| \le \sum_{\beta=1}^{d} \sup_t |\tilde{M}_n^{\beta,-}(t) - \tilde{W}_n^{\beta,-}(t)| \to_p 0.$$

Without loss we assume $\tilde{W}_n^-, \tilde{M}_n^-$ are scalar-valued processes and show

$$\sup_{0\le t\le 1} \big|\tilde{M}_n^-(t) - \tilde{W}_n^-(t)\big| \to_p 0.$$

Let $\epsilon > 0$ and choose k large enough so that $\|m - m^{(k)}\|_2 = \|\tilde{m} - \tilde{m}^{(k)}\|_2 < \epsilon$ and $\sum_{j=k}^{\infty} \|P^j v\|_1 < \epsilon^2$. Note that

$$\Big\|\sup_{0\le t\le 1} \big|\tilde{M}_n^-(t) - \tilde{W}_n^-(t)\big|\Big\|_2 \le \Big\|\sup_t \big|\tilde{M}_n^-(t) - \tilde{M}_n^{(k),-}(t)\big|\Big\|_2 + \Big\|\sup_t \big|\tilde{W}_n^-(t) - \tilde{M}_n^{(k),-}(t)\big|\Big\|_2.$$

Since $\tilde{W}_n^-(t), \tilde{M}_n^-(t), \tilde{M}_n^{(k),-}(t)$ are constant on each interval $[\frac{j}{n}, \frac{j+1}{n})$ for $j = 0, \ldots, n-1$, the supremum is attained at one of the points $t = 0, \frac{1}{n}, \ldots, \frac{n-1}{n}, 1$. Hence $\sup_{0\le t\le 1} \big|\sum_{j=1}^{[nt]} \tilde{v}\circ \tilde{T}^{-j}\big| = \max_{1\le \ell\le n} \big|\sum_{j=1}^{\ell} \tilde{v}\circ \tilde{T}^{-j}\big|$ and

$$\Big\|\sup_{0\le t\le 1} |\tilde{W}_n^-(t) - \tilde{M}_n^-(t)|\Big\|_2 \le n^{-1/2} \Big\|\overline{(\tilde{m} - \tilde{m}^{(k)})_n^-}\Big\|_2 + n^{-1/2} \Big\|\overline{(\tilde{v} - \tilde{m}^{(k)})_n^-}\Big\|_2. \tag{3.10}$$

We deal with the two terms in (3.10) separately. For the first term, note that $(\tilde{m} - \tilde{m}^{(k)}) \circ \tilde{T}^{-j}$ is a martingale difference sequence by Proposition 3.3.7. Hence apply Doob's inequality (Proposition 3.3.9) and use the orthogonality of $(\tilde{m} - \tilde{m}^{(k)}) \circ \tilde{T}^{-j}$

(Proposition 3.1.3) to calculate that

$$n^{-1/2}\left\|\overline{(\tilde{m}-\tilde{m}^{(k)})_n^-}\right\|_2 \le 2n^{-1/2}\left\|\sum_{j=1}^{n}(\tilde{m}-\tilde{m}^{(k)})\circ\tilde{T}^{-j}\right\|_2 = 2\|\tilde{m}-\tilde{m}^{(k)}\|_2 < 2\epsilon. \tag{3.11}$$

Next we consider the second term of (3.10). Use the martingale coboundary decomposition $v = m^{(k)} + \chi_1^k \circ T - \chi_1^k + P^k v$ to calculate that

$$\begin{aligned}
&n^{-1/2}\left\|\max_{1\le\ell\le n}\left|\sum_{j=1}^{\ell}(\tilde{v}-\tilde{m}^{(k)})\circ\tilde{T}^{-j}\right|\right\|_2 = n^{-1/2}\left\|\max_{\ell}\left|\sum_{j=1}^{\ell}(\chi_1^k\circ T-\chi_1^k+P^k v)\circ\pi\circ\tilde{T}^{-j}\right|\right\|_2\\
&= n^{-1/2}\left\|\max_{\ell}\left|\sum_{j=1}^{\ell}(\chi_1^k\circ\pi\circ\tilde{T}-\chi_1^k\circ\pi+P^k v\circ\pi)\circ\tilde{T}^{-j}\right|\right\|_2\\
&= n^{-1/2}\left\|\max_{\ell}\left|\chi_1^k\circ\pi\circ\tilde{T}^{-\ell+1}-\chi_1^k\circ\pi\circ\tilde{T}^{-1}+\sum_{j=1}^{\ell}(P^k v)\circ\pi\circ\tilde{T}^{-j}\right|\right\|_2\\
&\le 2n^{-1/2}\|\chi_1^k\|_\infty + n^{-1/2}\left\|\max_{\ell}\left|\sum_{j=1}^{\ell}(P^k v)\circ\pi\circ\tilde{T}^{-j}\right|\right\|_2\\
&\le 2n^{-1/2}k\|v\|_\infty + n^{-1/2}\left\|\max_{\ell}\left|\sum_{j=1}^{\ell}(P^k v)\circ\pi\circ\tilde{T}^{-j}\right|\right\|_2.
\end{aligned} \tag{3.12}$$

Next apply Rio's inequality (Proposition 3.3.10) to the square of the second term in (3.12) to see

$$\begin{aligned}
n^{-1}\left\|\max_{1\le\ell\le n}\left|\sum_{j=1}^{\ell}(P^k v)\circ\pi\circ\tilde{T}^{-j}\right|\right\|_2^2 &= n^{-1}\mathbb{E}\left(\max_{1\le\ell\le n}\left|\sum_{j=1}^{\ell}(P^k v)\circ\pi\circ\tilde{T}^{-j}\right|^2\right)\\
&\le 16n^{-1}\sum_{j=1}^{n}\max_{j\le\ell\le n} b_{j,\ell},
\end{aligned} \tag{3.13}$$

where

$$b_{j,\ell} = \left\|(P^k v)\circ\pi\circ\tilde{T}^{-j}\sum_{u=j}^{\ell}\mathbb{E}\big((P^k v)\circ\pi\circ\tilde{T}^{-u}|\mathcal{A}_j\big)\right\|_1.$$

Note that the observable $P^k v$ has a martingale coboundary decomposition

$$P^k v = r^{(k)} + \left(\sum_{n=1}^{\infty} P^n P^k v\right) \circ T - \sum_{n=1}^{\infty} P^n P^k v = r^{(k)} + \chi_{k+1}^{\infty} \circ T - \chi_{k+1}^{\infty},$$

with $r^{(k)} \in \operatorname{Ker} P,\ r^{(k)} \in L^2(\Lambda, \mathbb{R}^d)$ and $\chi_{k+1}^{\infty} \in L^1(\Lambda, \mathbb{R}^d)$. The proofs of these facts are identical to the ones proven for m in Theorem 3.1.2. Moreover, Proposition 3.3.7 shows $\tilde{r}^{(k)} \circ \tilde{T}^{-j} := r^{(k)} \circ \pi \circ \tilde{T}^{-j}$ forms a martingale difference sequence. That is, $\mathbb{E}(\tilde{r}^{(k)} \circ \tilde{T}^{-u} \,|\, \mathcal{A}_j) = 0$ for every $u > j$, and $\mathbb{E}(\tilde{r}^{(k)} \circ \tilde{T}^{-j} \,|\, \mathcal{A}_j) = \tilde{r}^{(k)} \circ \tilde{T}^{-j}$. Calculate that

$$\begin{aligned} b_{j,\ell} &\leq \|v\|_\infty \left\| \sum_{u=j}^{\ell} \mathbb{E}\Big(\big(r^{(k)} + \chi_{k+1}^{\infty} \circ T - \chi_{k+1}^{\infty}\big) \circ \pi \circ \tilde{T}^{-u} \,\Big|\, \mathcal{A}_j \Big) \right\|_1 \\ &= \|v\|_\infty \left\| r^{(k)} \circ \pi \circ \tilde{T}^{-j} + \mathbb{E}\left(\sum_{u=j}^{\ell} \chi_{k+1}^{\infty} \circ \pi \circ \tilde{T}^{-u+1} - \chi_{k+1}^{\infty} \circ \pi \circ \tilde{T}^{-u} \,\middle|\, \mathcal{A}_j \right) \right\|_1 \\ &\leq \|v\|_\infty \left(\big\|r^{(k)}\big\|_1 + 2\big\|\chi_{k+1}^{\infty}\big\|_1 \right). \end{aligned}$$

Returning to (3.13) we have

$$\begin{aligned} &n^{-1} \left\| \max_{1 \leq \ell \leq n} \left| \sum_{j=1}^{\ell} (P^k v) \circ \pi \circ \tilde{T}^{-j} \right| \right\|_2^2 \ll n^{-1} \sum_{j=1}^{n} \max_{j \leq \ell \leq n} \left(\|r^{(k)}\|_1 + 2\|\chi_{k+1}^{\infty}\|_1 \right) \\ &= \|P^k v + \chi_{k+1}^{\infty} \circ T - \chi_{k+1}^{\infty}\|_1 + 2\|\chi_{k+1}^{\infty}\|_1 \leq \|P^k v\|_1 + 4\|\chi_{k+1}^{\infty}\|_1 \leq 4 \sum_{n=k}^{\infty} \|P^n v\|_1 \ll \epsilon^2. \end{aligned}$$

Therefore (3.12) becomes

$$n^{-1/2} \left\| \overline{(\tilde{v} - \tilde{m}^{(k)})_n^-} \right\|_2 \ll n^{-1/2} k + \epsilon.$$

Combining this and (3.11) with (3.10), we see that

$$\limsup_{n \to \infty} \left\| \sup_{0 \leq t \leq 1} \left| \tilde{W}_n^-(t) - \tilde{M}_n^-(t) \right| \right\|_2 \ll \limsup_{n \to \infty} \left(2\epsilon + n^{-1/2} k + \epsilon \right) = 3\epsilon.$$

□

Let $\widehat{D}([0,1], \mathbb{R}^d)$ denote the space of $\mathbb{R}^d$-valued functions that are left continuous with right limits. This is the space of caǧlàd (not cadlaǧ) functions. This space and the following Proposition is needed because of our time reversal.

Proposition 3.3.12 ([29, Proposition 4.9]). *Suppose that* $A_n = B_n + F_n$*, where* $A_n \in D([0,1],\mathbb{R}^d)$*,* $B_n \in \widehat{D}([0,1],\mathbb{R}^d)$ *and* $\sup_{0\le t\le 1}|F_n(t)| \to_p 0$*. If* $B_n \to_w Z$ *in* $\big(\widehat{D}([0,1],\mathbb{R}^d),\mathcal{J}_1\big)$ *and* Z *has continuous sample paths, then* $A_n \to_w Z$ *in* $\big(D([0,1],\mathbb{R}^d),\|\cdot\|_\infty\big)$*.* □

The time reversal argument in the following proof is borrowed heavily from [29].

Proof of Theorem 3.3.3. Defining the map

$$g : (D([0,1],\mathbb{R}^d),\|\cdot\|_\infty) \to (\widehat{D}([0,1],\mathbb{R}^d),\|\cdot\|_\infty), \quad g(u)(t) := u(1) - u(1-t),$$

we see g is continuous since $\|g(u)-g(v)\|_\infty \le 2\|u-v\|_\infty$. Therefore Proposition 3.3.11 and the continuous mapping theorem (Proposition 3.2.9) shows $g(\tilde{W}_n^-) \to_w g(W)$ in $\big(\widehat{D}([0,1],\mathbb{R}^d),\mathcal{J}_1\big)$, where W is a d dimensional Brownian motion with mean zero and covariance $\Sigma = \int m\, m^T\, d\mu$. By Propositions 3.2.13 and 3.2.14, $\Sigma = \int v\, v^T\, d\mu + \sum_{j=1}^{\infty} \int v\,(v\circ T^j)^T + (v\circ T^j)\, v^T\, d\mu$. The result follows from Proposition 3.3.12 and the following two claims:

1. $\tilde{W}_n(t)\circ\tilde{T}^{-n} = g(\tilde{W}_n^-)(t) + F_n(t)$, where $\sup_{0\le t\le 1}|F_n(t)| \to_p 0$
2. $g(W) =_d W$

Assume the claims are true and denote $A_n = \tilde{W}_n(t)\circ\tilde{T}^{-n}$, $B_n = g(\tilde{W}_n^-)(t)$. We have $B_n \to_w g(W) =_d W$ in $\big(\widehat{D}([0,1],\mathbb{R}^d),\mathcal{J}_1\big)$, and $\sup_t|F_n(t)| \to_p 0$. Since W has continuous sample paths we see by Proposition 3.3.12 that $W_n =_d \tilde{W}_n =_d \tilde{W}_n\circ T^{-n} \to_w W$ in $\big(D([0,1],\mathbb{R}^d),\|\cdot\|_\infty\big)$. It remains to verify the two claims.

To prove the first claim, consider

$$\begin{aligned}\tilde{W}_n(t)\circ\tilde{T}^{-n} &= n^{-1/2}\sum_{j=0}^{[nt]-1}\tilde{v}\circ\tilde{T}^j\circ\tilde{T}^{-n} = n^{-1/2}\sum_{j=-n}^{[nt]-1-n}\tilde{v}\circ\tilde{T}^j = n^{-1/2}\sum_{j=n-[nt]+1}^{n}\tilde{v}\circ\tilde{T}^{-j}\\ &= n^{-1/2}\sum_{j=1}^{n}\tilde{v}\circ\tilde{T}^{-j} - n^{-1/2}\sum_{j=1}^{n-[nt]}\tilde{v}\circ\tilde{T}^{-j} = \tilde{W}_n^-(1) - n^{-1/2}\sum_{j=1}^{n-[nt]}\tilde{v}\circ\tilde{T}^{-j}.\end{aligned}$$

Define $F_n(t) := n^{-1/2}\big(\sum_{j=1}^{[n-nt]} - \sum_{j=1}^{n-[nt]}\big)\tilde{v}\circ\tilde{T}^{-j}$. Then

$$
\begin{aligned}
\tilde{W}_n(t)\circ\tilde{T}^{-n} &= \tilde{W}_n^-(1) - n^{-1/2}\sum_{j=1}^{n-[nt]}\tilde{v}\circ\tilde{T}^{-j} \\
&= \tilde{W}_n^-(1) + n^{-1/2}\left(\sum_{j=1}^{[n-nt]} - \sum_{j=1}^{[n-nt]} - \sum_{j=1}^{n-[nt]}\right)\tilde{v}\circ\tilde{T}^{-j} \\
&= \tilde{W}_n^-(1) - n^{-1/2}\sum_{j=1}^{[n-nt]}\tilde{v}\circ\tilde{T}^{-j} + F_n(t) \\
&= \tilde{W}_n^-(1) - \tilde{W}_n^-(1-t) + F_n(t) = g(\tilde{W}_n^-)(t) + F_n(t).
\end{aligned}
$$

If nt is an integer, then $F_n(t) = 0$. Otherwise $[n-nt] = n-[nt]-1$ and $F_n(t) = -n^{-1/2}\tilde{v}\circ\tilde{T}^{[nt]-n}$. It follows from Proposition 2.2.3 that $\sup_{0\le t\le 1}|F_n(t)| = n^{-1/2}\max_{0\le j\le n}|\tilde{v}\circ\tilde{T}^{-j}| \to_p 0$.

Finally, for the second claim, we show $g(W)(t) = W(1) - W(1-t)$ is a Brownian motion with mean zero and covariance Σ. Clearly $g(W)(0) = 0$. Next, $g(W)$ has continuous sample paths since W has continuous sample paths and g is a continuous function. Let $0 \le t_0 \le \cdots \le t_k \le 1$. Independent increments of $g(W)$ follows from independent increments of W, since

$$
\begin{aligned}
&\big(g(W)(t_1) - g(W)(t_0), \ldots, g(W)(t_k) - g(W)(t_{k-1})\big) \\
&= \big(W(1-t_0) - W(1-t_1), \ldots, W(1-t_{k-1}) - W(1-t_k)\big).
\end{aligned}
$$

Finally, note that for $0 \le s \le t \le 1$ we have

$$
g(W)(t) - g(W)(s) = W(1-s) - W(1-t) =_d \mathcal{N}(0, \Sigma(t-s)).
$$

Therefore $g(W) =_d W$.

□

3.4 Iterated Weak Invariance Principle

Recall the standing assumptions that $T : \Lambda \to \Lambda$ is an ergodic, measure-preserving map over the probability space $(\Lambda, \mathcal{A}, \mu)$, and our observable $v \in L^\infty(\Lambda, \mathbb{R}^d)$ is mean zero and satisfies the mixing assumption $\sum_{n=1}^\infty \|P^n v^i\|_1 < \infty$ for every $i = 1, \ldots, d$.

For coordinates $\beta, \gamma = 1, \ldots, d$, define the process

$$\mathbb{W}_n : [0,1] \to \mathbb{R}^{d\times d}, \quad \mathbb{W}_n^{\beta\gamma}(t) := n^{-1} \sum_{j=0}^{[nt]-1} \sum_{i=0}^{j-1} v^\beta \circ T^i \, v^\gamma \circ T^j.$$

Definition 3.4.1 (Itô integral, [31])**.** Let $X, Y \in D([0,1], \mathbb{R})$. For $0 < t \leq 1$, let Δ_ℓ be a sequence of refining partitions of $[0,t]$ such that $\max_{\Delta_\ell} |t_{k+1} - t_k| \to 0$. We define the *Itô integral* to be the following limit of Riemann-Stieltjes like sums, whenever it exists:

$$\int_0^t X \, dY = \lim_{\ell\to\infty} \sum_{k=0}^{\ell-1} X(t_k) \left(Y(t_{k+1}) - Y(t_k)\right),$$

where the limit is taken in probability.

Remark. If Y has paths of finite variation on compact subsets of $[0,1]$, then the Itô integral coincides with the Lebesgue-Stieltjes integral almost surely [46, Theorem 2.17].

Suppose X and Y are n dimensional vector-valued processes such that $\int X^\beta \, dY^\gamma$ exists, for every $\beta, \gamma = 1, \ldots, n$. Define the $n \times n$ dimensional matrix-valued process $\int X \otimes dY$ as

$$\left(\int X \otimes dY\right)^{\beta\gamma} := \int X^\beta \, dY^\gamma,$$

for $\beta, \gamma = 1, \ldots, n$.

Theorem 3.4.2 (Iterated WIP)**.** *We have that*

$$(W_n, \mathbb{W}_n) \to_w (W, \mathbb{W})$$

in $\left(D([0,1], \mathbb{R}^d \times \mathbb{R}^{d\times d}), \|\cdot\|_\infty\right)$*, where W is a d dimensional Brownian motion with mean zero and covariance* $\Sigma = \int v\, v^T \, d\mu + \sum_{r=1}^\infty \int v\,(v\circ T^r)^T + (v \circ T^r)\, v^T \, d\mu$*, and*

$$\mathbb{W}(t) = \int_0^t W \otimes dW + t \sum_{r=1}^\infty \int v \left(v \circ T^r\right)^T d\mu,$$

for $t \in [0,1]$.

Remark 1. As in Theorem 3.2.12, $\det \Sigma = 0$ if and only if there exists a non-zero $c \in \mathbb{R}^d$ such that $c^T v$ is an $L^1(\Lambda)$ coboundary.

Remark 2. We could equally define our processes in $D([0,K], \mathbb{R}^d)$ for any integer

$K \geq 1$, and the result still holds. Hence the result is true if we define our processes in $D([0,\infty),\mathbb{R}^d)$ by [7, Section 16].

Remark 3. If we replace Assumption (A) with the stronger condition $\sum_{n=1}^{\infty} \|P^n v\|_2 < \infty$, then the iterated WIP is already known by Kelly and Melbourne [29, Theorem 4.3].

Remark 4. If T is modelled by a Young tower [52, 53], then the iterated WIP is already known under Assumption (A) by Kelly and Melbourne [29, Theorem 10.2].

Remarks 3 and 4 tell us that Theorem 3.4.2 gives new examples for maps which are not modelled by a Young tower and that mix sufficiently slowly. Time-one maps of slowly mixing semi-flows are a rich source of such examples.

3.4.1 Suspensions and Examples

Suppose $T_t : M \to M$ is a semi-flow over some metric space (M,d) and that $v, w : M \to \mathbb{R}$ are some observables. For $0 < \eta \leq 1$ define $|v|_{C^\eta} := \sup_{x\neq x'} |v(x) - v(x')|/d(x,x')^\eta$ and

$$C^\eta(M) := \{v : M \to \mathbb{R} \,:\, \|v\|_{C^\eta} := |v|_{C^\eta} + \|v\|_\infty < \infty\}.$$

Let $|v|_{C^{0,\eta}} := \sup_{x\in M, t>0} |v(T_t x) - v(x)|/t^\eta$ and define the set of observables

$$C^{0,\eta}(M) := \{v : M \to \mathbb{R} \,:\, \|v\|_{C^\eta} + |v|_{C^{0,\eta}} < \infty\}.$$

Such observables are Hölder, and *Hölder in the flow direction.* Further define the *derivative in the flow direction* $\partial_t w := \lim_{t\to 0}(w \circ T_t - w)/t$, and let

$$L^{\infty,m}(M) := \{w : M \to \mathbb{R} \,:\, \|w\|_{L^{\infty,m}} := \sum_{j=0}^{m} \|\partial_t^j w\|_\infty < \infty\}$$

be the set of observables that have m derivatives in the flow direction that are $L^\infty(M)$.

The following example outlines a generic situation where the semi-flow attains the same mixing rate as a suitable Poincaré map.

Example 3.4.3. Suppose $T_t : M \to M$ is a semi-flow over an ambient, bounded metric space (M,d). Let $\Lambda \subseteq M$ be a Borel subset and suppose there exists a Hölder function $h : \Lambda \to \mathbb{R}^+$, with $\inf h > 0$, such that the map

$$Tx := T_{h(x)}(x) : \Lambda \to \Lambda$$

is modelled by a Young tower with polynomial tails [53]. For example Λ can be a Poincaré section for the semi-flow. Further assume that for any $C_1 > 0$ there exists $C_2 > 0$ such that

$$d(T_t x, T_t x') \leq C_2 d(x, x')^\eta \text{ for all } t \in [0, C_1] \text{ and } x, x' \in M. \tag{3.14}$$

Finally, assume absence of approximate eigenfunctions [37, Definition 3.4]. [37, Section 5] gives three different sufficient conditions for absence of approximate eigenfunctions. The first condition is imposed on periodic data of the semi-flow, and the set of roof functions that satisfy this condition is a prevalent [24] set. Prevalence is an analogue of almost-everywhere suitable for infinite dimensional spaces, and so absence of approximate eigenfunctions is satisfied with 'probability one'. The second condition is C^∞-dense and C^2-open; i.e. if there exists approximate eigenfunctions, we can make a C^∞ perturbation to the roof function h to ensure absence of approximate eigenfunctions. Let $v \in C^{0,\eta}(M)$ be a mean zero observable. Then [37, Theorem 3.11] guarantees a $\beta > 0$, $m \geq 1$ and $C > 0$ such that

$$\left| \int v\, w \circ T_t \, d\mu^h \right| \leq C\big(\|v\|_{C^\eta} + |v|_{C^{0,\eta}}\big)\|w\|_{L^{\infty,m}} t^{-\beta},$$

for all $w \in L^{\infty,m}(M)$. The rate β depends only on the polynomial tail of the the Young tower. The iterated WIP now holds for the time one map T_1 when $\beta > 1$. This gives new results for $\beta \in (1, 2]$. The case $\beta > 2$ follows from Kelly and Melbourne [29, Theorem 4.3].

We now outline a method of constructing semi-flows from a discrete map $T : \Lambda \to \Lambda$, and give an example where the mixing rate of T passes over to the semi-flow.

Assume (Λ, d) is a metric space and consider an integrable function $h : [0, 1] \to \mathbb{R}^+$. Define the space $\Lambda^h := \big(\Lambda \times \mathbb{R}^+\big)/\sim$, where we make the identification $(x, h(x)) \sim (Tx, 0)$. We call h a *roof function* and define the *suspension semi-flow* T_t over h as

$$\begin{aligned} &T_t : \Lambda^h \to \Lambda^h \\ &T_t(x, u) = (x, u + t) \bmod \sim \end{aligned} \tag{3.15}$$

Figure 3.3 depicts a graphical representation of T_t. Defining $\bar{h} := \int h \, d\mu$ we have that $\mu^h := \mu \times \text{Leb}/\bar{h}$ is a probability measure on Λ^h.

Proposition 3.4.4. *If T is measure-preserving with respect to μ on Λ, then T_t is measure-preserving with respect to μ^h on Λ^h.*

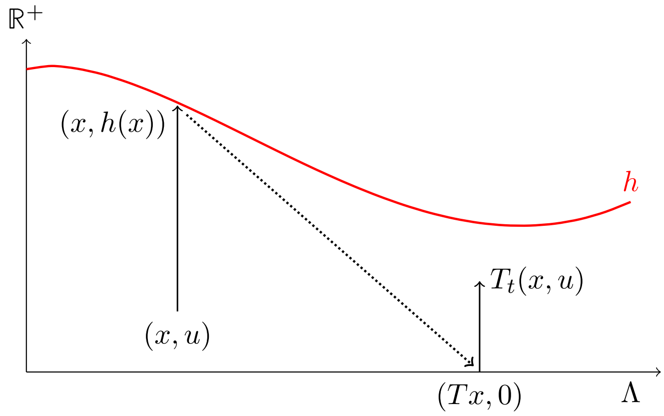

Figure 3.3: A graphical representation of the suspension flow T_t.

Proof. Let $f \in L^1(\Lambda^h)$ and calculate that

$$\begin{aligned}
\int f \circ T_t \, d\mu^h &= \bar{h}^{-1} \int_0^{h(x)} \int_\Lambda f(T_t(x,s)) \; d\mu(x) ds = \bar{h}^{-1} \int_0^{h(x)} \int_\Lambda f(x, s+t) \; d\mu(x) ds \\
&= \bar{h}^{-1} \int_t^{h(x)+t} \int_\Lambda f(x,s) \; d\mu(x) ds \\
&= \bar{h}^{-1} \left(\int_0^{h(x)} + \int_{h(x)}^{h(x)+t} - \int_0^t \right) \int_\Lambda f(x,s) \; d\mu(x) ds \\
&= \int f \, d\mu^h + \bar{h}^{-1} \left(\int_{h(x)}^{h(x)+t} - \int_0^t \right) \int_\Lambda f(x,s) \; d\mu(x) ds.
\end{aligned}$$

We see that the second term is zero since

$$\int_{h(x)}^{h(x)+t} \int_\Lambda f(x,s) \; d\mu(x) ds = \int_0^t \int_\Lambda f(Tx, s) \; d\mu(x) ds = \int_0^t \int_\Lambda f(x,s) \; d\mu(x) ds,$$

and the result follows. $\square$

Example 3.4.5. Let $0 < \alpha < 1$ and let T be the intermittent map introduced in Example 3.2.2:

$$Tx = \begin{cases} x(1 + 2^\alpha x^\alpha) & \text{if } 0 \leq x \leq 1/2 \\ 2x - 1 & \text{if } 1/2 < x \leq 1 \end{cases}.$$

Recall that for mean zero Hölder observables $v \in C^\eta([0,1])$ we have $\|P^n v\|_1 \leq C n^{-1/\alpha+1}$, and so Assumption (A) holds when $0 < \alpha < 1/2$. Since T is modelled by a Young tower, the iterated WIP already holds by [29]. To obtain new examples, we lift to a suitable semi-flow and consider the time-one map.

Let $h : [0,1] \to \mathbb{R}^+$ be an η-Hölder roof function such that $\inf h > 0$ and let $T_t : \Lambda^h \to \Lambda^h$ be the corresponding suspension flow. Further assume absence of approximate eigenfunctions for T_t and assume condition (3.14) holds with respect to the 'metric' induced on Λ^h, $d\big((x,u),(y,t)\big) := d(x,y) + |u-t|$. Note that d is not really a metric on Λ^h, since $d\big((x,h(x)),(Tx,0)\big) = d(x,Tx)+h(x)$. Let $v \in C^{0,\eta}(\Lambda^h)$ be a mean zero observable. Then [37, Theorem 3.11] guarantees an $m \geq 1$ and $C > 0$ such that

$$\left|\int v\, w \circ T_t \, d\mu^h\right| \leq C\big(\|v\|_{C^\eta} + |v|_{C^{0,\eta}}\big)\|w\|_{L^{\infty,m}} t^{-1/\alpha+1},$$

for all $w \in L^{\infty,m}(\Lambda^h)$. Proposition 2.1.3 shows assumption (A) holds for the time-one map T_1 when $0 < \alpha < 1/2$, and therefore Theorem 3.4.2 gives the iterated WIP for T_1. Since T_1 is not modelled by a Young tower, this gives new results in the range $1/3 \leq \alpha < 1/2$.

Remark. Any (semi-)flow with a neutral direction different to the flow direction cannot be modelled by a Young tower.

3.4.2 Proof of the Iterated WIP

Define

$$\mathbb{X}_n : [0,1] \to \mathbb{R}^{d\times d}, \quad \mathbb{X}_n^{\beta\gamma}(t) := n^{-1} \sum_{j=0}^{[nt]-1} \sum_{i=0}^{j-1} m^\beta \circ T^i \, v^\gamma \circ T^j.$$

Proposition 3.4.6 (Martingale iterated WIP)**.** *We have that*

$$(W_n, \mathbb{X}_n) \to_w (W, \mathbb{X})$$

in $\big(D([0,1], \mathbb{R}^d \times \mathbb{R}^{d\times d}), \|\cdot\|_\infty\big)$*, where W is a d dimensional Brownian motion with mean zero and covariance* $\Sigma = \int m\, m^T \, d\mu$*, and*

$$\mathbb{X}(t) = \int_0^t W \otimes dW$$

for $t \in [0,1]$*.*

The proof of this proposition is postponed until the next subsection. We first prove the iterated WIP:

Proof of Theorem 3.4.2. Let $\Gamma^{\beta\gamma}(t) := t\sum_{r=1}^\infty \int v^\beta\, v^\gamma \circ T^r \, d\mu$. Note that $(W_n, \mathbb{W}_n) = (W_n, \mathbb{X}_n) + (0, \mathbb{W}_n - \mathbb{X}_n)$. By Proposition 3.4.6, $(W_n, \mathbb{X}_n) \to_w (W, \mathbb{X})$ with $\Sigma = \int m\, m^T \, d\mu$. Propositions 3.2.13 and 3.2.14 give the correct covariance. By Proposition 3.3.1 it remains to show $\mathbb{W}_n - \mathbb{X}_n \to_p \Gamma$. This is equivalent to the component-wise statement:

$$\sup_{0\le t\le 1}\left|\mathbb{W}_n^{\beta\gamma}(t)-\mathbb{X}_n^{\beta\gamma}(t)-t\sum_{r=1}^{\infty}\int v^\beta\, v^\gamma\circ T^r\,d\mu\right|\to_p 0,$$

for each $\beta,\gamma=1,\dots,d$. Instead of convergence in probability, we prove convergence almost surely.

By Proposition 2.2.2 it suffices to show

$$\mathbb{W}_n^{\beta\gamma}(1)-\mathbb{X}_n^{\beta\gamma}(1)\to_{a.s}\sum_{j=1}^{\infty}\int v^\beta\, v^\gamma\circ T^j\,d\mu,$$

which follows from the following calculation:

$$\mathbb{W}_n^{\beta\gamma}(1)-\mathbb{X}_n^{\beta\gamma}(1)=n^{-1}\sum_{j=0}^{n-1}\sum_{i=0}^{j-1}\Big(v^\beta\circ T^i\, v^\gamma\circ T^j-m^\beta\circ T^i\, v^\gamma\circ T^j\Big)$$
$$=n^{-1}\sum_{j=0}^{n-1}\sum_{i=0}^{j-1}\big(\chi^\beta\circ T-\chi^\beta\big)\circ T^i\, v^\gamma\circ T^j=n^{-1}\sum_{j=0}^{n-1}(\chi^\beta\circ T^j-\chi^\beta)\, v^\gamma\circ T^j$$
$$=n^{-1}\sum_{j=0}^{n-1}(\chi^\beta v^\gamma)\circ T^j-\chi^\beta n^{-1}\sum_{j=0}^{n-1}v^\gamma\circ T^j\to_{a.s.}\int\chi^\beta\, v^\gamma\,d\mu=\sum_{r=1}^{\infty}\int v^\beta\, v^\gamma\circ T^r\,d\mu.$$

□

Proof of Proposition 3.4.6

Here we follow a technique similar to one used in [29]. We first introduce the lifted space to consider the backwards processes.

By Proposition 3.2.4, lift the probability space $(\Lambda,\mathcal{A},\mu)$ and map T to obtain a probability space $(\tilde\Lambda,\tilde{\mathcal{A}},\tilde\mu)$, where $\tilde\mu$ is an ergodic, invariant measure with respect to the invertible map $\tilde T:\tilde\Lambda\to\tilde\Lambda$. Associated to this is a measure-preserving projection $\pi:\tilde\Lambda\to\Lambda$ such that $\pi\circ\tilde T=T\circ\pi$. Define the lifted observables $\tilde v:=v\circ\pi:\tilde\Lambda\to\mathbb{R}^d$ and $\tilde m=m\circ\pi$.

Next define the lifted forwards process $\tilde W_n$ the same as the forward processes W_n but with v,T replaced with $\tilde v,\tilde T$ respectively. Define the lifted backwards process

$$\tilde W_n^-:[0,1]\to\mathbb{R}^d,\quad \tilde W_n^-(t)=n^{-1/2}\sum_{j=1}^{[nt]}\tilde v\circ\tilde T^{-j},$$

and define $\tilde M_n^-(t)=n^{-1/2}\sum_{j=1}^{[nt]}\tilde m\circ\tilde T^{-j}$. Recall that $\sum_1^0\tilde v\circ\tilde T^{-j}=\sum_1^0\tilde m\circ\tilde T^{-j}=0$. We define our iterated processes as follows. For coordinates $\beta,\gamma=1,\dots,d$ define the lifted process

$$\tilde{\mathbb{X}}_n : [0,1] \to \mathbb{R}^{d\times d}, \quad \tilde{\mathbb{X}}_n^{\beta\gamma}(t) := n^{-1} \sum_{j=0}^{[nt]-1} \sum_{i=0}^{j-1} \tilde{m}^\beta \circ \tilde{T}^i \, \tilde{v}^\gamma \circ \tilde{T}^j,$$

and the lifted backwards process

$$\tilde{\mathbb{X}}_n^- : [0,1] \to \mathbb{R}^{d\times d}, \quad \tilde{\mathbb{X}}_n^{\beta\gamma,-}(t) := n^{-1} \sum_{j=1}^{[nt]-1} \sum_{i=j+1}^{[nt]} \tilde{v}^\beta \circ \tilde{T}^{-j} \, \tilde{m}^\gamma \circ \tilde{T}^{-i}.$$

To prove Proposition 3.4.6 we verify the hypothesis of Proposition 3.3.12 with $A_n = (W_n, \mathbb{X}_n)$ and $B_n = \chi(\tilde{W}_n^-, \tilde{M}_n^-, \tilde{\mathbb{X}}_n^-)$ for some continuous function χ. The remainder of this section is as follows: First, we verify that $A_n = B_n + F_n$, where $\sup_{0\le t\le 1} |F_n(t)| \to_p 0$. Next, we show $(\tilde{W}_n^-, \tilde{M}_n^-, \tilde{\mathbb{X}}_n^-) \to_w (W, W, \mathbb{X})$. Finally, we show $B_n = \chi(\tilde{W}_n^-, \tilde{M}_n^-, \tilde{\mathbb{X}}_n^-) \to_w \chi(W, W, \mathbb{X}) =: Z$ and $Z =_d (W, \mathbb{X})$.

Verifying $A_n = B_n + F_n$

Define

$$g : D([0,1], \mathbb{R}^d) \to \widehat{D}([0,1], \mathbb{R}^d), \quad g(u)(t) := u(1) - u(1-t),$$
$$h : D([0,1], \mathbb{R}^d) \times D([0,1], \mathbb{R}^d) \to \widehat{D}([0,1], \mathbb{R}^{d\times d}), \quad h(u,v)(t) := u(1-t)\big(v(1) - v(1-t)\big)^T.$$

Proposition 3.4.7. *We have that*

$$\big(W_n, \mathbb{X}_n\big) =_d \left(g(\tilde{W}_n^-), \Big(g(\tilde{\mathbb{X}}_n^-) - h(\tilde{W}_n^-, \tilde{M}_n^-) \Big)^T \right) + F_n,$$

where F_n is such that $\sup_{0\le t\le 1} |F_n(t)| \to_p 0$.

Proof. Since π and $\tilde{T}$ are measure-preserving, we see

$$\big(W_n, \mathbb{X}_n\big) =_d \big(\tilde{W}_n, \tilde{\mathbb{X}}_n\big) \circ \tilde{T}^{-n}.$$

Hence it suffices to show

$$\big(\tilde{W}_n, \tilde{\mathbb{X}}_n\big) \circ \tilde{T}^{-n} = \left(g(\tilde{W}_n^-), \Big(g(\tilde{\mathbb{X}}_n^-) - h(\tilde{W}_n^-, \tilde{M}_n^-) \Big)^T \right) + F_n, \tag{3.16}$$

where F_n is to be defined and $\sup_{0\le t\le 1} |F_n(t)| \to_p 0$. We now drop the tilde notation.

The first coordinate of (3.16) has already been shown in the proof of Theorem 3.3.3. Since convergence in probability can be shown coordinatewise, it remains to show

$$\mathbb{X}_n^{\beta\gamma}(t)\circ T^{-n} = g(\mathbb{X}_n^-)^{\gamma\beta}(t) - h(W_n^-, M_n^-)^{\gamma\beta}(t) + F_n(t),$$

and $\sup_{0\le t\le 1}|F_n(t)| \to_p 0$. Consider

$$\begin{aligned}
\mathbb{X}_n^{\beta\gamma}(t)\circ T^{-n} &= n^{-1}\sum_{j=0}^{[nt]-1}\sum_{i=0}^{j-1} m^\beta\circ T^{i-n}\, v^\gamma\circ T^{j-n} = n^{-1}\sum_{j=-n}^{[nt]-n-1}\sum_{i=0}^{j+n-1} m^\beta\circ T^{i-n}\, v^\gamma\circ T^{j} \\
&= n^{-1}\sum_{j=-n}^{[nt]-n-1}\sum_{i=-n}^{j-1} m^\beta\circ T^{i}\, v^\gamma\circ T^{j} = n^{-1}\sum_{j=n-[nt]+1}^{n}\sum_{i=-n}^{-j-1} m^\beta\circ T^{i}\, v^\gamma\circ T^{-j} \\
&= n^{-1}\sum_{j=n-[nt]+1}^{n}\sum_{i=j+1}^{n} m^\beta\circ T^{-i}\, v^\gamma\circ T^{-j} \\
&= n^{-1}\left(\sum_{j=1}^{n-1}+\sum_{j=n}^{n}-\sum_{j=1}^{[n-nt]-1}-\sum_{j=[n-nt]}^{n-[nt]}\right)\sum_{i=j+1}^{n} m^\beta\circ T^{-i}\, v^\gamma\circ T^{-j} \\
&= \mathbb{X}_n^{\gamma\beta,-}(1) + F_n^1(t) - A_n(t) - F_n^2(t), \qquad (3.17)
\end{aligned}$$

where

$$\begin{aligned}
F_n^1(t) &= n^{-1}v^\gamma\circ T^{-n}\sum_{i=n+1}^{n} m^\beta\circ T^{-i} = \left(n^{-1}v^\gamma\circ T^{-n}\right)\left(\sum_{i=1}^{0} m^\beta\circ T^{-i}\right)\circ T^{-n} = 0, \\
A_n(t) &= n^{-1}\sum_{j=1}^{[n-nt]-1}\sum_{i=j+1}^{n} m^\beta\circ T^{-i}\, v^\gamma\circ T^{-j}, \\
F_n^2(t) &= n^{-1}\sum_{j=[n-nt]}^{n-[nt]}\sum_{i=j+1}^{n} m^\beta\circ T^{-i}\, v^\gamma\circ T^{-j}.
\end{aligned}$$

Next, calculate that

$$
\begin{aligned}
A_n(t) &= n^{-1} \sum_{j=1}^{[n-nt]-1} \sum_{i=j+1}^{n} m^\beta \circ T^{-i}\, v^\gamma \circ T^{-j} \\
&= n^{-1} \sum_{j=1}^{[n-nt]-1} \left(\sum_{i=j+1}^{[n-nt]} + \sum_{i=[n-nt]+1}^{n} \right) m^\beta \circ T^{-i}\, v^\gamma \circ T^{-j} \\
&= \mathbb{X}_n^{\gamma\beta,-}(1-t) \\
&\qquad + \left(n^{-1/2} \sum_{i=[n-nt]+1}^{n} m^\beta \circ T^{-i} \right) \left(n^{-1/2} \sum_{j=1}^{[n-nt]} v^\gamma \circ T^{-j} - n^{-1/2} v^\gamma \circ T^{-[n-nt]} \right) \\
&= \mathbb{X}_n^{\gamma\beta,-}(1-t) + \Big(M_n^{\beta,-}(1) - M_n^{\beta,-}(1-t) \Big) W_n^{\gamma,-}(1-t) - F_n^3(t),
\end{aligned}
$$

where

$$
F_n^3(t) = n^{-1/2} v^\gamma \circ T^{-[n-nt]} \Big(M_n^{\beta,-}(1) - M_n^{\beta,-}(1-t) \Big) .
$$

Combining equations (3.17) and (**??**), we have

$$
\begin{aligned}
\mathbb{X}_n^{\beta\gamma}(t) \circ T^{-n} &= \mathbb{X}_n^{\gamma\beta,-}(1) - A_n(t) - F_n^2(t) \\
&= \mathbb{X}_n^{\gamma\beta,-}(1) - \mathbb{X}_n^{\gamma\beta,-}(1-t) - W_n^{\gamma,-}(1-t) \Big(M_n^{\beta,-}(1) - M_n^{\beta,-}(1-t) \Big) - F_n^2(t) + F_n^3(t) \\
&= g(\mathbb{X}_n^-)^{\gamma\beta}(t) - h(W_n^-, M_n^-)^{\gamma\beta}(t) - F_n^2(t) + F_n^3(t).
\end{aligned}
$$

Finally, we show $\sup_t |F_n^2(t)|$, $\sup_t |F_n^3(t)| \to_p 0$. For the first part, since $[n-nt] \geq$

$n-[nt]-1$ we see

$$
\begin{aligned}
|F_n^2(t)| &= n^{-1}\left|\sum_{j=[n-nt]}^{n-[nt]}\sum_{i=j+1}^{n} m^\beta\circ T^{-i}\,v^\gamma\circ T^{-j}\right| \\
&\le n^{-1}\sum_{j=n-[nt]-1}^{n-[nt]}\left|\sum_{i=j+1}^{n} m^\beta\circ T^{-i}\,v^\gamma\circ T^{-j}\right| \\
&= n^{-1}\left|v^\gamma\circ T^{[nt]-n+1}\sum_{i=n-[nt]}^{n} m^\beta\circ T^{-i}\right| \\
&\qquad + n^{-1}\left|v^\gamma\circ T^{[nt]-n}\sum_{i=n-[nt]+1}^{n} m^\beta\circ T^{-i}\right| \\
&\le \left(n^{-1/2}\left|v^\gamma\circ T^{[nt]-n+1}\right|\right)\left(n^{-1/2}\left|m^\beta\circ T^{-(n-[nt])}\right| + \left|M_n^\beta(t)\circ T^{-n}\right|\right) \\
&\qquad + \left(n^{-1/2}\left|v^\gamma\circ T^{[nt]-n}\right|\right)\left(\left|M_n^\beta(t)\circ T^{-n}\right|\right).
\end{aligned}
$$

Hence

$$
\begin{aligned}
\sup_{0\le t\le 1}|F_n^2(t)| &\le n^{-1/2}\max_{0\le j\le n}|v^\gamma|\circ T^{-j+1}\left(n^{-1/2}\max_{0\le j\le n}|m^\beta|\circ T^{-j} + \sup_{0\le t\le 1}\left|M_n^\beta(t)\circ T^{-n}\right|\right) \\
&\qquad + n^{-1/2}\max_{0\le j\le n}|v^\gamma|\circ T^{-j}\sup_{0\le t\le 1}\left|M_n^\beta(t)\circ T^{-n}\right| \\
&\le 2n^{-1/2}\max_{0\le j\le n}|v^\gamma|\circ T^{-j}\sup_{0\le t\le 1}|M_n^\beta(t)\circ T^{-n}| \\
&\qquad + n^{-1}\max_{0\le j\le n}|m^\beta\,v^\gamma\circ T|\circ T^{-j}. \qquad (3.18)
\end{aligned}
$$

For F_n^3, calculate that

$$
\begin{aligned}
\sup_{0\le t\le 1}|F_n^3(t)| &= \sup_{0\le t\le 1}\left|n^{-1/2}v^\gamma\circ T^{-[n-nt]}\left(M_n^{\beta,-}(1)-M_n^{\beta,-}(1-t)\right)\right| \\
&\le 2n^{-1/2}\max_{1\le j\le n}|v^\gamma|\circ T^{-j}\sup_{0\le t\le 1}|M_n^{\beta,-}(t)|. \qquad (3.19)
\end{aligned}
$$

Since $v^\gamma\in L^2(\Lambda)$ and $m^\beta v^\gamma\circ T\in L^1(\Lambda)$, Proposition 2.2.3 shows

$$
n^{-1/2}\max_{1\le j<n}|v^\gamma|\circ T^{-j}\to_{a.s} 0, \qquad n^{-1}\max_{1\le j\le n}|m^\beta\,v^\gamma\circ T|\circ T^{-j}\to_{a.s} 0.
$$

Therefore the second term of (3.18) converges to zero almost surely. Finally, we show (3.19) and the first term of (3.18) converges to zero in probability. Note that by Proposition 3.3.8 we have $M_n^{\beta,-}\to_w W^\beta$, and so the continuous mapping theorem shows $\sup_t |M_n^{\beta,-}(t)|,\ \sup_t |M_n^{\beta}(t)| \to_w \sup_t |W^{\beta}(t)|$. In particular, $\left(\sup_t |M_n^{\beta,-}(t)|\right)_{n\geq 1}$, $\left(\sup_t |M_n^{\beta}(t)|\right)_{n\geq 1}$ are tight sequences. Proposition 2.2.4 now shows

$$n^{-1/2} \max_{1\leq j\leq n} |v^\gamma| \circ T^{-j} \sup_{0\leq t\leq 1} |M_n^{\beta,-}(t)| \to_p 0,$$
$$n^{-1/2} \max_{1\leq j\leq n} |v^\gamma| \circ T^{-j} \sup_{0\leq t\leq 1} |M_n^{\beta}(t) \circ T^{-n}| \to_p 0,$$

and therefore $\sup_t |F_n^2(t)|,\ \sup_t |F_n^3(t)| \to_p 0$. □

Showing $(\tilde{W}_n^-, \tilde{M}_n^-, \tilde{\mathbb{X}}_n^-) \to_w (W, W, \mathbb{X})$

To show $(\tilde{W}_n^-, \tilde{M}_n^-, \tilde{\mathbb{X}}_n^-) \to_w (W, W, \mathbb{X})$, we need a suitable convergence result:

Proposition 3.4.8 ([26], [31, Theorem 2.2])**.** *For each $n \geq 1$, let (X_n, Y_n) be an $\mathcal{A}_{n,t}$-measurable process with sample paths in $D([0,1], \mathbb{R}^d \times \mathbb{R}^d)$ and let Y_n be a martingale with respect to $\mathcal{A}_{n,t}$. Suppose $\sup_n \mathbb{E}([Y_n^\beta, Y_n^\beta](t)) < \infty$ for each $\beta = 1, \ldots, d$ and $t \in [0,1]$, and $(X_n, Y_n) \to_w (X, Y)$ in $\big(D([0,1], \mathbb{R}^d \times \mathbb{R}^d), \|\cdot\|_\infty\big)$. Then*

$$\left(X_n, Y_n, \int X_n \otimes dY_n\right) \to_w \left(X, Y, \int X \otimes dY\right)$$

in $\big(D([0,1], \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}), \|\cdot\|_\infty\big)$. □

Proposition 3.4.9. *For $\beta, \gamma = 1, \ldots, d$ and $t \in [0,1]$ we have that the process $\tilde{\mathbb{X}}_n^-$ is given by the following Riemann-Stieltjes integral:*

$$\tilde{\mathbb{X}}_n^-(t) = \int_0^t \tilde{W}_n^- \otimes d\tilde{M}_n^-.$$

Proof. In this proof we drop the tilde notation. Consider a sequence of partitions Δ_ℓ of $[0,t]$ such that $\max_{\Delta_\ell} |t_{k+1} - t_k| \to 0$ as $\ell \to \infty$. By definition,

$$\int_0^t W_n^{\beta,-}\, dM_n^{\gamma,-} = \lim_{\ell\to\infty} S_\ell,$$

where

$$S_\ell = \sum_{k=0}^{\ell-1} W_n^{\beta,-}(t_k)\big(M_n^{\gamma,-}(t_{k+1}) - M_n^{\gamma,-}(t_k)\big).$$

Take ℓ large enough so that $\max_{k<\ell} |t_{k+1} - t_k| < \frac{1}{n}$. For $j = 0, 1, \ldots, [nt] - 1$, the map $s \mapsto \sum_{j=1}^{[ns]} v \circ T^{-j}$ is constant on each interval $[\frac{j}{n}, \frac{j+1}{n})$, and so S_ℓ depends only on each jump v_j, for $j = 0, 1, \ldots, [nt]-1$. Thus

$$\begin{aligned} S_\ell &= n^{-1} \sum_{j=0}^{[nt]-1} v_j^{\beta,-} \big(m_{j+1}^{\gamma,-} - m_j^{\gamma,-}\big) = n^{-1} \sum_{j=0}^{[nt]-1} \left(\sum_{i=1}^{j} v^\beta \circ T^{-i} \right) m^\gamma \circ T^{-(j+1)} \\ &= n^{-1} \sum_{j=1}^{[nt]} \sum_{i=1}^{j-1} v^\beta \circ T^{-i} m^\gamma \circ T^{-j} = n^{-1} \sum_{1 \le i < j \le [nt]} v^\beta \circ T^{-i} m^\gamma \circ T^{-j}. \end{aligned}$$

This can be written as

$$\begin{aligned} n^{-1} \sum_{1 \le i < j \le [nt]} v^\beta \circ T^{-i} m^\gamma \circ T^{-j} &= n^{-1} \sum_{i=1}^{[nt]} \sum_{j=i+1}^{[nt]} v^\beta \circ T^{-i} m^\gamma \circ T^{-j} \\ &= n^{-1} \sum_{i=1}^{[nt]-1} \sum_{j=i+1}^{[nt]} v^\beta \circ T^{-i} m^\gamma \circ T^{-j} = \mathbb{X}_n^{\beta\gamma,-}(t), \end{aligned}$$

since taking $i = [nt]$ in the outer sum gives $n^{-1} v^\beta \circ T^{-[nt]} \sum_{j=[nt]+1}^{[nt]} m^\gamma \circ T^{-j} = 0$. Thus we obtain $\int_0^t W_n^- \otimes dM_n^- = \mathbb{X}_n^-(t)$ as required. $\square$

Proposition 3.4.10 (Backward iterated WIP)**.** *We have the following convergence:*

$$\big(\tilde{W}_n^-, \tilde{M}_n^-, \tilde{\mathbb{X}}_n^-\big) \to_w \big(W, W, \mathbb{X}\big)$$

in $\big(D([0,1], \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d \times d}), \|\cdot\|_\infty\big)$.

Proof. For this proof we drop the tilde notation. We verify the hypothesis of Proposition 3.4.8 with

$$X_n = W_n^-, \quad Y_n = M_n^-.$$

We see Y_n is a martingale by Proposition 3.3.7. The condition $\sup_n \mathbb{E}\big([Y_n^\beta, Y_n^\beta](t)\big) < \infty$ follows from Proposition 3.3.6 since

$$\mathbb{E}\big([M_n^{\beta,-}, M_n^{\beta,-}](t)\big) = \frac{1}{n} \sum_{j=1}^{[nt]} \int (m^\beta)^2 \circ T^{-j} \, d\mu = \frac{[nt]}{n} \int (m^\beta)^2 \, d\mu \le t \|m^\beta\|_2^2. \quad (3.20)$$

To verify weak convergence of (X_n, Y_n), first note that $M_n^- \to_w W$ by Corollary 3.3.8, which implies $(M_n^-, M_n^-) \to_w (W, W)$ by the continuous mapping theorem. Moreover, the proof of Proposition 3.3.11 shows $\sup_{0 \le t \le 1} |W_n^-(t) - M_n^-(t)| \to_p 0$, so $(W_n^-, M_n^-) \to_w (W, W)$. Proposition 3.4.8 now applies and we obtain the desired convergence

$$\left(W_n^-, M_n^-, \int W_n^- \otimes dM_n^-\right) \to_w \left(W, W, \int W \otimes dW\right),$$

in $\big(D([0,1], \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}), \|\cdot\|_\infty\big)$. $\square$

Checking $B_n \to_w Z$ and $Z =_d (W, \mathbb{X})$

Proposition 3.4.11. *We have that*

$$\left(g(\tilde{W}_n^-), \left(g(\tilde{\mathbb{X}}_n^-) - h(\tilde{W}_n^-, \tilde{M}_n^-)\right)^T\right) \to_w \left(g(W), \big(g(\mathbb{X}) - h(W,W)\big)^T\right)$$

in $\big(\widehat{D}([0,1], \mathbb{R}^d \times \mathbb{R}^{d\times d}), \|\cdot\|_\infty\big)$.

Proof. We easily obtain the desired convergence by Proposition 3.4.10 and the continuous mapping theorem once we establish continuity of g and h. It was shown in the proof of Theorem 3.3.3 that g is continuous. We see h is continuous since $h^{\beta\gamma}$ is the product of two continuous maps, for each $\beta, \gamma = 1, \dots, d$. $\square$

Proposition 3.4.12. *We have that*

$$\left(g(W), \big(g(\mathbb{X}) - h(W,W)\big)^T\right) =_d (W, \mathbb{X}).$$

Proof. Fix $t \in [0,1]$ and coordinates $\beta, \gamma = 1, \dots, d$. Let $\mathbb{G}^{\beta\gamma}(t) := \int_0^t g(W)^\beta \, dg(W)^\gamma$. We show $(W, \mathbb{X}) =_d (g(W), \mathbb{G})$ and $\mathbb{G} = \big(g(\mathbb{X}) - h(W,W)\big)^T$.

To see $(W, \mathbb{X}) =_d (g(W), \mathbb{G})$, note that

$$\mathbb{X}(t) = \lim_{n\to\infty} X_n(t), \quad \mathbb{G}(t) = \lim_{n\to\infty} Y_n(t),$$

where the limits are taken in probability and

$$X_n^{\beta\gamma}(t) = \sum_{k=0}^{[nt]-1} W^\beta\left(\frac{k}{n}\right)\left[W^\gamma\left(\frac{k+1}{n}\right) - W^\gamma\left(\frac{k}{n}\right)\right],$$

$$Y_n^{\beta\gamma}(t) = \sum_{k=0}^{[nt]-1} g(W)^\beta\left(\frac{k}{n}\right)\left[g(W)^\gamma\left(\frac{k+1}{n}\right) - g(W)^\gamma\left(\frac{k}{n}\right)\right].$$

Denote the identity map by I. Since the map

$$
\begin{aligned}
&f_n : D([0,1], \mathbb{R}^d) \to D([0,1], \mathbb{R}^{d\times d}),\\
&\big(f_n(X)\big)^{\beta\gamma}(t) = \sum_{k=0}^{[nt]-1} X^\beta\left(\frac{k}{n}\right)\left[X^\gamma\left(\frac{k+1}{n}\right) - X^\gamma\left(\frac{k}{n}\right)\right]
\end{aligned}
$$

is continuous, $W =_d g(W)$ and weak equality is preserved under continuous maps, we have

$$(W, X_n) = (I, f_n)W =_d (I, f_n)g(W) = (g(W), Y_n).$$

Moreover since $X_n \to_p \mathbb{X}$ and $Y_n \to_p \mathbb{G}$, we see

$$(W, X_n) \to_w (W, \mathbb{X}), \quad (g(W), Y_n) \to_w (g(W), \mathbb{G}).$$

By uniqueness of weak limits, $(W, \mathbb{X}) =_d (g(W), \mathbb{G})$.

It remains to show $\mathbb{G}^{\beta\gamma}(t) = g(\mathbb{X})^{\gamma\beta}(t) - h(W,W)^{\gamma\beta}(t)$. Recall that

$$\mathbb{G}^{\beta\gamma}(t) = \int_0^t g(W)^\beta \, dg(W)^\gamma = \lim_{n\to\infty} Y_n^{\beta\gamma}(t),$$

where the limit is taken in probability and

$$
\begin{aligned}
Y_n^{\beta\gamma}(t) &= \sum_{k=0}^{[nt]-1} g(W)^\beta\left(\frac{k}{n}\right)\left[g(W)^\gamma\left(\frac{k+1}{n}\right) - g(W)^\gamma\left(\frac{k}{n}\right)\right]\\
&= \sum_{k=0}^{[nt]-1}\left[W^\beta(1) - W^\beta\left(1-\frac{k}{n}\right)\right]\left[W^\gamma\left(1-\frac{k}{n}\right) - W^\gamma\left(1-\frac{k+1}{n}\right)\right]\\
&= \sum_{k=0}^{[nt]-1}\left[\sum_{j=0}^{k-1} W^\beta\left(1-\frac{j}{n}\right) - W^\beta\left(1-\frac{j+1}{n}\right)\right]\left[W^\gamma\left(1-\frac{k}{n}\right) - W^\gamma\left(1-\frac{k+1}{n}\right)\right]\\
&= \sum_{j=0}^{[nt]-1}\sum_{k=j+1}^{[nt]-1}\left[W^\beta\left(1-\frac{j}{n}\right) - W^\beta\left(1-\frac{j+1}{n}\right)\right]\left[W^\gamma\left(1-\frac{k}{n}\right) - W^\gamma\left(1-\frac{k+1}{n}\right)\right]\\
&= \sum_{j=0}^{[nt]-1}\left[W^\beta\left(1-\frac{j}{n}\right) - W^\beta\left(1-\frac{j+1}{n}\right)\right]\left[W^\gamma\left(1-\frac{j+1}{n}\right) - W^\gamma\left(1-\frac{[nt]}{n}\right)\right].
\end{aligned}
$$

Now consider

$$\big(g(\mathbb{X}) - h(W,W)\big)^{\gamma\beta}(t) = \mathbb{X}^{\gamma\beta}(1) - \mathbb{X}^{\gamma\beta}(1-t) - W^{\gamma}(1-t)\left(W^{\beta}(1) - W^{\beta}(1-t)\right)$$
$$= \int_0^1 W^{\gamma}\, dW^{\beta} - \int_0^{1-t} W^{\gamma}\, dW^{\beta} - W^{\gamma}(1-t)\left(\int_0^1 dW^{\beta} - \int_0^{1-t} dW^{\beta}\right)$$
$$= \int_{1-t}^1 \big(W^{\gamma} - W^{\gamma}(1-t)\big)\, dW^{\beta} = \lim_{n\to\infty} Z_n^{\beta\gamma}(t),$$

where the limit is taken in probability and

$$Z_n^{\beta\gamma}(t) = \sum_{k=[n-nt]}^{n-1} \left[W^{\gamma}\left(\frac{k}{n}\right) - W^{\gamma}(1-t)\right]\left[W^{\beta}\left(\frac{k+1}{n}\right) - W^{\beta}\left(\frac{k}{n}\right)\right].$$

With the change of coordinates $j = n-k-1$ we see

$$Z_n^{\beta\gamma}(t) = \sum_{j=0}^{n-[n-nt]-1} \left[W^{\gamma}\left(1-\frac{j+1}{n}\right) - W^{\gamma}(1-t)\right]\left[W^{\beta}\left(1-\frac{j}{n}\right) - W^{\beta}\left(1-\frac{j+1}{n}\right)\right]. \quad (3.21)$$

By uniqueness of limits in probability, the result will follow once we establish that $Z_n - Y_n \to_p 0$. We instead show the stronger statement $\sup_{0\le t\le 1}|Z_n^{\beta\gamma}(t) - Y_n^{\beta\gamma}(t)| \to_{a.s.} 0$. First suppose nt is not an integer, so $[n-nt] = n-[nt]-1$ and the upper limit in the sum in equation (3.21) is now $n-[n-nt]-1 = [nt]$. Denoting $W_j^{\beta} := W^{\beta}\big(1-\frac{j}{n}\big) - W^{\beta}\big(1-\frac{j+1}{n}\big)$ we have

$$Z_n^{\beta\gamma}(t) - Y_n^{\beta\gamma}(t) = \sum_{j=0}^{[nt]} W_j^{\beta}\left[W^{\gamma}\left(1-\frac{j+1}{n}\right) - W^{\gamma}(1-t)\right]$$
$$- \sum_{j=0}^{[nt]-1} W_j^{\beta}\left[W^{\gamma}\left(1-\frac{j+1}{n}\right) - W^{\gamma}\left(1-\frac{[nt]}{n}\right)\right]$$
$$= \sum_{j=0}^{[nt]-1} W_j^{\beta}\left[W^{\gamma}\left(1-\frac{[nt]}{n}\right) - W^{\gamma}(1-t)\right]$$
$$+ \sum_{j=[nt]}^{[nt]} W_j^{\beta}\left[W^{\gamma}\left(1-\frac{j+1}{n}\right) - W^{\gamma}(1-t)\right]$$
$$= A_n(t) + B_n(t),$$

where

$$
\begin{aligned}
A_n(t) &= \sum_{j=0}^{[nt]-1} W_j^{\beta} \left[W^{\gamma}\left(1 - \frac{[nt]}{n}\right) - W^{\gamma}(1-t) \right] \\
&= \left[W^{\beta}(1) - W^{\beta}\left(1 - \frac{[nt]}{n}\right) \right] \left[W^{\gamma}\left(1 - \frac{[nt]}{n}\right) - W^{\gamma}(1-t) \right], \\
B_n(t) &= \sum_{j=[nt]}^{[nt]} W_j^{\beta} \left[W^{\gamma}\left(1 - \frac{j+1}{n}\right) - W^{\gamma}(1-t) \right] \\
&= \left[W^{\beta}\left(1 - \tfrac{[nt]}{n}\right) - W^{\beta}\left(1 - \tfrac{[nt]+1}{n}\right) \right] \left[W^{\gamma}\left(1 - \tfrac{[nt]+1}{n}\right) - W^{\gamma}(1-t) \right].
\end{aligned}
$$

Since $\dfrac{[nt]}{n}, \dfrac{[nt]+1}{n} \to t$ and W is almost surely uniformly continuous, we have $\sup_t |A_n(t)|, \sup_t |B_n(t)| \to_{a.s.} 0$ and hence $\sup_t |Z_n^{\beta\gamma}(t) - Y_n^{\beta\gamma}(t)| \to_{a.s} 0$. When nt is an integer, the upper limit in the sum in equation (3.21) is $n-[n-nt]-1 = [nt]-1$, so $B_n = 0$ and the above convergence still holds. $\square$

Proof of Proposition 3.4.6

Proof of Proposition 3.4.6. We verify the hypothesis of Proposition 3.3.12 with

$$
A_n = (W_n, \mathbb{X}_n), \quad B_n = \left(g(\tilde{W}_n^-), \left(g(\tilde{\mathbb{X}}_n^-) - h(\tilde{W}_n^-, \tilde{M}_n^-) \right)^T \right).
$$

Proposition 3.4.7 shows $A_n = B_n + F_n$ and $\sup_{0 \le t \le 1} |F_n(t)| \to_p 0$. Proposition 3.4.11 shows $B_n \to_w Z$ for some process Z and Proposition 3.4.12 shows $Z =_d (W, \mathbb{X})$. Therefore, Proposition 3.3.12 shows

$$
(W_n, \mathbb{X}_n) \to_w (W, \mathbb{X}),
$$

in $\left(D([0,1], \mathbb{R}^d \times \mathbb{R}^{d \times d}), \|\cdot\|_\infty\right)$ as desired. $\square$

Chapter 4

Limit Theorems for Invertible Maps

For this chapter, suppose $T : \Lambda \to \Lambda$ is an invertible, ergodic and measure-preserving map over the probability space $(\Lambda, \mathcal{A}, \mu)$ and let $v : \Lambda \to \mathbb{R}$ be a mean zero $L^\infty(\Lambda)$ observable.

We do not need to lift our map T to obtain an invertible one, and hence the backwards Birkhoff sums $v_n^- = \sum_{j=1}^n v \circ T^{-j}$ are always defined. We denote the n^{th} backwards Birkhoff sum with a subscript n and a minus superscript. Unfortunately, however, the transfer operator is $Pv = v \circ T^{-1}$ (see the proof of Proposition 2.1.2). As a consequence of T being measure-preserving, the mixing condition $\sum_{n=1}^\infty \|P^n v\|_1 < \infty$ from Section 3 can never be non-trivially satisfied. To remedy this, we suppose there exists a sub-sigma algebra $\mathcal{A}_0 \subset \mathcal{A}$ such that $T^{-1}\mathcal{A}_0 \subseteq \mathcal{A}_0$. Then $\mathbb{E}(v \circ T^{-1} \,|\, \mathcal{A}_0)$ plays the role of the transfer operator. Many authors [11, 18, 32, 42] assume existence of such a sigma algebra.

If there exists a partition $\mathcal{W}^s = \{W^s(x) : x \in \Lambda\}$ of Λ—where $W^s(x)$ is a (measurable) set containing x—such that $TW^s(x) \subseteq W^s(Tx)$ for all $x \in \Lambda$, then we say $\mathcal{W}^s$ is a *stable foliation* for the map T and we call $W^s(x)$ a *local stable leaf.* We say T *contracts stable leaves exponentially* if there exists $C > 0$ and $\gamma \in (0,1)$ such that for every $x \in \Lambda$ and $x' \in W^s(x)$ we have $d(T^n x', T^n x) \le C\gamma^n$. The sigma algebra generated by the partition, $\mathcal{A}_0 := \sigma(\mathcal{W}^s)$, is a sub-sigma algebra such that $T^{-1}\mathcal{A}_0 \subseteq \mathcal{A}_0$.

Let $L^p(\Lambda, \mathcal{A}_0) := \{v \in L^p(\Lambda) : v \text{ is } \mathcal{A}_0\text{-measurable}\}$. For $n \in \mathbb{Z}$, define $\mathcal{A}_n := T^{-n}\mathcal{A}_0$ and $\mathbb{E}_n v := \mathbb{E}(v \,|\, \mathcal{A}_n)$. For any $\ell, p \in \mathbb{Z}$ we have

$$\mathbb{E}_\ell \mathbb{E}_p v = \mathbb{E}\big(\mathbb{E}(v \,|\, T^{-\ell}\mathcal{A}_0) \,\big|\, T^{-p}\mathcal{A}_0\big) = \mathbb{E}(v \,|\, T^{-\max\{\ell,p\}}\mathcal{A}_0) = \mathbb{E}_{\max\{\ell,p\}} v,$$

and

$$(\mathbb{E}_\ell v)\circ T^p = \mathbb{E}(v\,|\,T^{-\ell}\mathcal{A}_0)\circ T^p = \mathbb{E}(v\circ T^p\,|\,T^{-\ell-p}\mathcal{A}_0) = \mathbb{E}_{\ell+p}(v\circ T^p).$$

A standard probabilistic assumption is

$$\sum_{n=0}^{\infty}\|\mathbb{E}_0(v\circ T^n)-v\circ T^n\|_1+\sum_{n=1}^{\infty}\big\|\mathbb{E}_0(v\circ T^{-n})\big\|_1<\infty, \tag{A$_{\text{inv}}$}$$

and we assume this holds throughout the following chapter, unless stated otherwise.

4.0.1 Sufficient Criteria for Assumption (A$_{\text{inv}}$)

We do not assume (A$_{\text{inv}}$) in this subsection.

Proposition 4.0.1 ([12]). *Assume there exists $\beta>1$ and $C>0$ such that for all $n\geq 1$,*

$$\left|\int v\,w\circ T^n\,d\mu\right|\leq C\|w\|_\infty n^{-\beta}\ \textit{for all}\ w\in L^\infty(\Lambda,\mathcal{A}_0).$$

Then

$$\|\mathbb{E}_0(v\circ T^{-n})\|_1\leq C\|v\|_\infty n^{-\beta}.$$

Proof. Define $\zeta:=\operatorname{sgn}\big(\mathbb{E}_n v\big)$. Clearly ζ is $\mathcal{A}_n$-measurable. Using the hypothesis and the fact that $\|E_0(v\circ T^{-n})\|_1=\|\mathbb{E}_n v\|_1$ we see

$$\begin{aligned}\|\mathbb{E}_n v\|_1&=\int\mathbb{E}_n v\,\zeta\,d\mu=\int\mathbb{E}_n(v\,\zeta)\,d\mu=\int v\,\zeta\,d\mu=\int v\ \operatorname{sgn}(\mathbb{E}_n v)\circ T^{-n+n}\,d\mu\\&=\int v\ \operatorname{sgn}\big(\mathbb{E}_0(v\circ T^{-n})\big)\circ T^n\,d\mu\leq C\|v\|_\infty n^{-\beta}.\end{aligned}$$

□

Proposition 4.0.2 ([12]). *Assume that (Λ,d) is a metric space. The observable $v:\Lambda\to\mathbb{R}$ only needs to be mean zero and measurable here. Suppose T admits a stable foliation $\mathcal{W}^s$ and let $Y\subseteq\Lambda$ be a positive measure subset that is a union of stable leaves in $\mathcal{W}^s$. Define the first return time $\tau:Y\to\mathbb{N}$, $\tau(y):=\inf\{n\geq 1\,:\,T^n y\in Y\}$ and induced map $F:=T^\tau:Y\to Y$. Let $h_n(x):=\#\{0\leq j\leq n\,:\,T^j x\in Y\}$. Suppose that $\mu(\tau>n)\leq Cn^{-\alpha}$ for some $\alpha>1$, and that there are constants $C\geq 1$, $\gamma\in(0,1)$ such that*

$$\Big|\operatorname{diam}\big(v(T^n W^s)\big)\Big|\leq C\gamma^{h_n(x)}\ \ \textit{for all}\ W^s\in\mathcal{W}^s,\ n\geq 1. \tag{4.1}$$

Then

$$\|\mathbb{E}_0(v \circ T^n) - v \circ T^n\|_1 \leq Cn^{-\alpha+1}.$$

□

Remark. Suppose $S : \Lambda \to \Lambda$ is another invertible, ergodic and measure-preserving map that admits the same stable foliation as T. The conclusion of Proposition 4.0.2 remains valid when equation (4.1) is replaced with

$$\left|\operatorname{diam}\big(v(S^n W^s)\big)\right| \leq C\gamma^{h_n(x)} \quad \text{for all } W^s \in \mathcal{W}^s,\ n \geq 1, \tag{4.1$'$}$$

where h_n remains unchanged.

The following definition is a subset of the original definition of Young [52], but it suffices for our purposes.

Definition 4.0.3 (Young tower, [40, Page 3])**.** We say T *is modelled by a Young tower with tail* τ if there is a subset Y and a countable partition Y_j of $Y \subseteq \Lambda$ such that τ is constant on each Y_j. There is a measurable partition $\mathcal{W}^s$ of Y (stable leaves) and each Y_j is a union of elements of $\mathcal{W}^s$. Let W^u denote a measurable subset that intersects each stable leaf only once. The induced map $F = T^\tau : Y \to Y$ satisfies $F(W^s(y)) \subset W^s(Fy)$, and we define the separation time $s(y, y')$ as the least integer $n \geq 0$ such that $F^n y, F^n y'$ lie in distinct partition elements. Further suppose there exists constants $C \geq 1$, $\gamma \in (0, 1)$ such that

$$\begin{aligned} d(T^\ell F^j y,\, T^\ell F^j y') &\leq C\gamma^j \text{ for all } y' \in W^s(y), y \in Y, \\ d(T^\ell F^j y,\, T^\ell F^j y') &\leq C\gamma^{s(y,y')-j} \text{ for all } y, y' \in W^u, \end{aligned} \tag{4.2}$$

for all $j \geq 0$, $0 \leq \ell \leq r(F^j y)$. Let $\bar{Y} = Y/\sim$, where $y \sim y'$ if $y' \in W^s(y)$. Since $F(W^s(y)) \subset W^s(Fy)$, we obtain a well defined quotient map $\bar{F} : \bar{Y} \to \bar{Y}$. Assume $\bar{F}$ is Gibbs-Markov [1, Chapter 4] and assume $\tau \in L^1(Y)$.

Proposition 4.0.4 ([12])**.** *Assume* (Λ, d) *is a metric space and* $v : \Lambda \to \mathbb{R}$ *is a Hölder continuous and mean zero observable. If* T *is modelled by a Young tower with return time* τ*, and tail rate* $\mu(\tau > n) \leq Cn^{-\beta}$ *for some* $C > 0$ *and* $\beta > 1$*, then*

$$\|\mathbb{E}_0(v \circ T^{-n})\|_1 \leq Cn^{-\beta+1}, \quad \|\mathbb{E}_0(v \circ T^n) - v \circ T^n\|_1 \leq Cn^{-\beta+1}.$$

□

4.0.2 An Example – Intermittent Baker's map

For $0 < e, k < 1$, let $T_{ek} : \Lambda \to \Lambda$, $\Lambda := [0,1]^2$ be the intermittent Baker's map

$$T_{ek}(x,y) = \begin{cases} (f_e(x), f_k^{-1}(y)), & (x,y) \in [0,1/2) \times [0,1] \\ (2x-1,(y+1)/2), & (x,y) \in [1/2,1] \times [0,1] \end{cases}, \tag{4.3}$$

where $f_e(x) = x(1+2^e x^e)$. Note that $T_{ek}(x,y) = \big(T_e^{(1)}x, T_k^{(2)}y\big)$ has a skew-product structure. The map $T_e^{(1)}$ is the intermittent map from Example 3.2.2 and sees only expansion, whereas the map $T_k^{(2)}$ sees only contraction. The stable leaves are vertical lines. A smaller parameter corresponds to stronger expansion/contraction (taking the parameter $e = 0$ in $T_e^{(1)}$ gives the doubling map). Let $v : \Lambda \to \mathbb{R}$ be a Hölder continuous, mean zero observable.

The case $k = e$ is considered by Melbourne and Varandas [40, Example 4.1], and following this example, there is an ergodic invariant measure μ for $T := T_{ek}$. Moreover defining $Y := [1/2,1] \times [0,1]$, T is modelled by a Young tower with return time map $\tau(\underline{x}) := \inf\{n \geq 1 : T^n\underline{x} \in Y\}$. Since Y extends fully in the y-direction, the return time τ is identical to return time of the intermittent map $T_e^{(1)}$ to $[1/2,1]$ and so [53, Proof of Theorem 5(c)] shows $\mu(\tau > n) \leq Cn^{-1/e}$. Theorem 4.0.4 now shows

$$\|\mathbb{E}_0(v \circ T^{-n})\|_1 \leq Cn^{-1/e+1}, \quad \|\mathbb{E}_0(v \circ T^n) - v \circ T^n\|_1 \leq Cn^{-1/e+1},$$

for mean zero, Hölder observables. Assumption ($\mathrm{A_{inv}}$) now holds for all $0 < e = k < 1/2$.

For the iterated WIP, it will be necessary to construct an example with stronger contraction than expansion:

Proposition 4.0.5. *Suppose $0 < k < e < 1/2$. Let T_{ek} be the map defined in (4.3), and let $v : \Lambda \to \mathbb{R}$ be a Hölder continuous, mean zero observable. Then we have*

$$\|\mathbb{E}_0(v \circ T^{-n})\|_1 \leq Cn^{-1/e+1}, \quad \|\mathbb{E}_0(v \circ T^n) - v \circ T^n\|_1 \leq Cn^{-1/k+1}.$$

Proof. Denote $T := T_{ek}$ and define $Y := [1/2,1] \times [0,1]$, $\tau_e(\underline{x}) := \inf\{n \geq 1 : T_{ee}^n\,\underline{x} \in Y\}$. We first show T is modelled by a Young tower with tail τ_e. Since T has a product structure and the stable leaves are the same for every choice of parameters $0 < k, e < 1/2$, the only condition we need to check is the exponential contraction on returns to Y (condition (4.2)). This follows from the claim that $\tau_k \leq \tau_e$ for every $k < e$. Assuming the claim, we have $\tau_e = \tau_k + \tilde{\tau}$ for the function

$\tilde{\tau} := \tau_e - \tau_k : Y \to \mathbb{N}_0$. Denoting $F := T^{\tau_e}$, we have for any $\underline{y} \in Y$, $\underline{y}' \in W^s(\underline{y})$, $j \geq 0$ and $0 \leq \ell \leq \tau_e(F^j \underline{y})$ that

$$d(T^\ell F^j \underline{y},\, T^\ell F^j \underline{y}') = d(T^\ell \big(T^{\tau_k + \tilde{\tau}}\big)^j \underline{y},\, T^\ell \big(T^{\tau_k + \tilde{\tau}}\big)^j \underline{y}') = d(T^{\ell + j\tilde{\tau}} \big(T^{\tau_k}\big)^j \underline{y},\, T^{\ell + j\tilde{\tau}} \big(T^{\tau_k}\big)^j \underline{y}')$$

From the product structure and the fact $T_{kk}^{\tau_k}$ contracts stable leaves exponentially, we see T^{τ_k} also contracts stable leaves exponentially. Further, since $\tilde{\tau}$ is constant on stables leaves, it follows that there exists a $p \geq 0$, depending on the value of $\ell + j\tilde{\tau}$, such that

$$d(T^{\ell + j\tilde{\tau}} \big(T^{\tau_k}\big)^j \underline{y},\ T^{\ell + j\tilde{\tau}} \big(T^{\tau_k}\big)^j \underline{y}') \leq C\gamma^{-j-p} \leq C\gamma^{-j},$$

for some $C > 1$, $\gamma \in (0,1)$, and so (4.2) is satisfied. Therefore T is modelled by a Young tower with tail τ_e and Theorem 4.0.4 shows

$$\|E_0(v \circ T^{-n})\|_1 \leq Cn^{-1/e+1}.$$

We now prove the claim: Denote $S_e := T_e^{(1)}$ with return time $R_e(x) := \inf\{n \geq 1 : S_e^n x \in [1/2, 1]\}$. Since Y extends fully in the y-direction and T has a product structure, $\tau_e = R_e$. Therefore it suffices to show that $R_k \leq R_e$ for $k < e$. Recall that

$$S_k x = \begin{cases} x(1 + 2^k x^k), & x \in [0, 1/2) \\ 2x - 1 & x \in [1/2, 1] \end{cases}.$$

Let $y \in [1/2, 1]$. Since the second branch does not depend on the parameter k, we see $S_k y = S_e y =: x_0$. If $x_0 \in [1/2, 1]$, then $R_k(y) = R_e(y) = 1$ and there is nothing to show. Suppose $x_0 \in [0, 1/2)$. We now want to show that the orbit of x_0 under S_k hits $[1/2, 1]$ before the orbit of x_0 under S_e does. Since S_k, S_e are increasing functions, it suffices to show that for any $x \in [0, 1/2)$ we have $S_k x \geq S_e x$. This follows easily since $2^k x^k \geq 2^e x^e$. Therefore $\tau_k \leq \tau_e$.

For the second condition, we show (4.1′) holds and apply Proposition 4.0.2. Note that for any $\underline{x}_0 \in [0,1] \times [0,1]$ we have

$$\begin{aligned} \left| \operatorname{diam}\Big(v\big(T^n W^s(\underline{x}_0)\big)\Big) \right| &= \sup_{\underline{x}, \underline{x}' \in W^s(\underline{x}_0)} \big| v(T^n \underline{x}) - v(T^n \underline{x}') \big| \\ &\leq |v|_{C^\eta} \sup_{\underline{x}, \underline{x}' \in W^s(\underline{x}_0)} d\big(T^n \underline{x}, T^n \underline{x}'\big) \end{aligned} \tag{4.4}$$

Due to the product structure, the contraction on stable leaves does not depend on the first coordinate of $T = T_{ek}$, so we can replace T on the right hand side with T_{kk}

to obtain

$$\left|\operatorname{diam}\left(v\big(T^n W^s(\underline{x}_0)\big)\right)\right| \le |v|_{C^\eta} \sup_{\underline{x},\underline{x}'\in W^s\underline{x}_0} d\big(T^n_{kk}\underline{x}, T^n_{kk}\underline{x}'\big) \tag{4.5}$$

Since $T_{kk}^{\tau_k}$ contracts stable leaves exponentially, we see (4.1′) holds and so

$$\|\mathbb{E}_0(v\circ T^n) - v\circ T^n\|_1 \le Cn^{-1/k+1}.$$

□

4.1 A Martingale Coboundary Decomposition

For $\ell > k \ge 0$ define the forwards part of the transfer function

$$\chi_k^\ell := \sum_{n=k}^{\ell} \mathbb{E}_0(v\circ T^n) - v\circ T^n,$$

and for $\ell > k \ge 1$ define the backwards part as

$$\chi_{-k}^{-\ell} := \sum_{n=-k}^{-\ell} \mathbb{E}_0(v\circ T^n).$$

Now, for $k < 0 < \ell$ we define $\chi_k^\ell := \chi_k^{-1} + \chi_0^\ell$. By ($\mathrm{A}_{\mathrm{inv}}$), $\chi := \chi_{-\infty}^{\infty}$ is well-defined and lies in $L^1(\Lambda)$.

Proposition 4.1.1. *We have the following martingale coboundary decomposition:*

$$v = m + \chi\circ T - \chi,$$

where $m \in L^1(\Lambda)$, m is $\mathcal{A}_0$-measurable, and $\mathbb{E}_1 m = 0$.

Proof. Denoting $V_n := v\circ T^n$, define

$$\begin{aligned}
m &:= V_0 + \chi - \chi\circ T = V_0 + \chi_0^\infty - \chi_0^\infty\circ T + \chi_{-\infty}^{-1} - \chi_{-\infty}^{-1}\circ T \\
&= V_0 + \sum_{n=0}^{\infty} \mathbb{E}_0 V_n - V_n - (\mathbb{E}_0 V_n)\circ T + V_{n+1} + \mathbb{E}_0 V_{-n-1} - (\mathbb{E}_0 V_{-n-1})\circ T \\
&= \sum_{n=0}^{\infty} \mathbb{E}_0 V_n - \mathbb{E}_1 V_{n+1} + \mathbb{E}_0 V_{-n-1} - \mathbb{E}_1 V_{-n}.
\end{aligned}$$

It is clear that $m \in L^1(\Lambda)$ since $v, \chi \in L^1(\Lambda)$. Since $\mathcal{A}_1 \subseteq \mathcal{A}_0$, we also see that m is

$\mathcal{A}_0$-measurable. It remains to show that $\mathbb{E}_1 m = 0$. This follows from the calculation

$$E_1 m = \mathbb{E}_1 \left(\sum_{n=0}^{\infty} \mathbb{E}_0 V_n - \mathbb{E}_1 V_{n+1} + \mathbb{E}_0 V_{-n-1} - \mathbb{E}_1 V_{-n} \right) = 0.$$

□

An important hypothesis in the martingale limit theorems that we apply is that the random variables are square integrable. The next two propositions introduce an 'intermediate' truncated decomposition, and then use this truncated decomposition to show $m \in L^2(\Lambda)$.

Proposition 4.1.2. *For each $k \geq 1$, v has the decomposition*

$$v = m^{(k)} + \chi_{-k}^k \circ T - \chi_{-k}^k + \mathbb{E}_0(v \circ T^{-k}) - \mathbb{E}_0(v \circ T^{k+1}) + v \circ T^{k+1},$$

where $m^{(k)} \in L^\infty(\Lambda)$, $m^{(k)}$ is $\mathcal{A}_0$-measurable, and $\mathbb{E}_1 m^{(k)} = 0$. Moreover, $m^{(k)} \to m$ in $L^1(\Lambda)$.

Proof. Denote $V_n := v \circ T^n$. For $k \geq 1$ define

$$\begin{aligned} m^{(k)} &:= V_0 + \chi_{-k}^k - \chi_{-k}^k \circ T - \mathbb{E}_0 V_{-k} + \mathbb{E}_0 V_{k+1} - V_{k+1} \\ &= V_0 - V_{k+1} + \sum_{n=0}^{k} (\mathbb{E}_0 V_n - V_n - (\mathbb{E}_0 V_n) \circ T + V_{n+1}) \\ &\qquad + \sum_{n=1}^{k} (\mathbb{E}_0 V_{-n} - (\mathbb{E}_0 V_{-n}) \circ T) + \mathbb{E}_0 V_{k+1} - \mathbb{E}_0 V_{-k} \\ &= \mathbb{E}_0 V_{k+1} - \mathbb{E}_0 V_{-k} + \sum_{n=0}^{k} (\mathbb{E}_0 V_n - \mathbb{E}_1 V_{n+1}) + \sum_{n=1}^{k} (\mathbb{E}_0 V_{-n} - \mathbb{E}_1 V_{-n+1}). \end{aligned} \tag{4.6}$$

It follows that $m^{(k)}$ is $\mathcal{A}_0$-measurable. To see $m^{(k)} \in L^\infty(\Lambda)$, use the fact that $\|\chi_{-k}^k\|_\infty \leq (3k+2)\|v\|_\infty$ to calculate

$$\|m^{(k)}\|_\infty \leq \|V_0 - \mathbb{E}_0 V_{-k} + \mathbb{E}_0 V_{k+1} - V_{k+1}\|_\infty + (6k+4)\|v\|_\infty \leq (6k+8)\|v\|_\infty < \infty.$$

Next, $\mathbb{E}_1 m^{(k)} = 0$ follows from the fact that $\mathbb{E}_1\mathbb{E}_0 = \mathbb{E}_1$ and (4.6), since

$$\begin{aligned}
\mathbb{E}_1 m^{(k)} &= \mathbb{E}_1\left(\mathbb{E}_0 V_{k+1} - \mathbb{E}_0 V_{-k} + \sum_{n=0}^{k}(\mathbb{E}_0 V_n - \mathbb{E}_1 V_{n+1}) + \sum_{n=1}^{k}(\mathbb{E}_1 V_{-n} - \mathbb{E}_1 V_{-n+1})\right) \\
&= \mathbb{E}_1 V_{k+1} - \mathbb{E}_1 V_{-k} + \sum_{n=0}^{k}(\mathbb{E}_1 V_n - \mathbb{E}_1 V_{n+1}) + \sum_{n=1}^{k}(\mathbb{E}_1 V_{-n} - \mathbb{E}_1 V_{-n+1}) \\
&= \mathbb{E}_1 V_{k+1} - \mathbb{E}_1 V_{-k} + \mathbb{E}_1 V_0 - \mathbb{E}_1 V_{k+1} + \mathbb{E}_1 V_{-k} - \mathbb{E}_1 V_0 = 0.
\end{aligned}$$

Finally, we see $m^{(k)} \to m$ in $L^1(\Lambda)$ by calculating

$$\begin{aligned}
\|m - m^{(k)}\|_1 &= \left\|\chi - \chi_{-k}^{k} - \chi \circ T + \chi_{-k}^{k} \circ T + \mathbb{E}_0(V_{-k}) - \big(\mathbb{E}_0(V_{k+1}) - V_{k+1}\big)\right\|_1 \\
&= \left\|\chi_{-\infty}^{-k-1} + \chi_{k+1}^{\infty} - \chi_{-\infty}^{-k-1} \circ T - \chi_{k+1}^{\infty} \circ T + \mathbb{E}_0(V_{-k}) - \big(\mathbb{E}_0(V_{k+1}) - V_{k+1}\big)\right\| \\
&\le \|\chi_{-\infty}^{-k} - \chi_{-\infty}^{-k-1} \circ T\|_1 + \|\chi_{k+2}^{\infty} - \chi_{k+1}^{\infty} \circ T\|_1 \le 2\|\chi_{-\infty}^{-k}\|_1 + 2\|\chi_{k}^{\infty}\|_1 \\
&\le 2\sum_{n=k}^{\infty}\big(\|E_0(v \circ T^{-n})\|_1 + \|\mathbb{E}_0(v \circ T^{n}) - v \circ T^{n}\|_1\big) \to 0.
\end{aligned}$$

□

Proposition 4.1.3. *We have $m \in L^2(\Lambda)$ and $m^{(k)} \to m$ in $L^2(\Lambda)$.*

Proof. By Proposition 4.1.2, $m^{(k)} \to m$ in $L^1(\Lambda)$. Therefore by uniqueness of limits in $L^1(\Lambda)$ it suffices to show that $m^{(k)}$ is a Cauchy sequence in $L^2(\Lambda)$. For clarity denote $V_n := v \circ T^n$ and

$$a_n = \begin{cases} \mathbb{E}_0 V_n - V_n, & n \ge 0 \\ \mathbb{E}_0 V_n, & n < 0 \end{cases}.$$

Assume without loss that $\ell > k > 0$. With this notation, we have

$$m^{(\ell)} = v - \chi_{-\ell}^{\ell} \circ T + \chi_{-\ell}^{\ell} - a_{-\ell} + a_{\ell+1},$$

and $\chi_{-k}^{k} = \sum_{j=-k}^{k} a_j$. Calculate that

$$\begin{aligned}
m^{(\ell)} - m^{(k)} &= \big(\chi_{-\ell}^{\ell} - \chi_{-k}^{k}\big) - \big(\chi_{-\ell}^{\ell} \circ T - \chi_{-k}^{k} \circ T\big) + a_{\ell+1} - a_{-\ell} - \big(a_{k+1} - a_{-k}\big) \\
&= \big(\chi_{-\ell}^{-k-1} + \chi_{k+1}^{\ell}\big) - \big(\chi_{-\ell}^{-k-1} \circ T + \chi_{k+1}^{\ell} \circ T\big) + a_{\ell+1} - a_{-\ell} - \big(a_{k+1} - a_{-k}\big) \\
&= \big(\chi_{-\ell}^{-k-1} - \chi_{-\ell}^{-k-1} \circ T - a_{-\ell} + a_{-k}\big) + \big(\chi_{k+1}^{\ell} - \chi_{k+1}^{\ell} \circ T + a_{\ell+1} - a_{k+1}\big) \\
&= \big(\chi_{-\ell+1}^{-k} - \chi_{-\ell}^{-k-1} \circ T\big) + \big(\chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T\big).
\end{aligned}$$

Hence

$$\|m^{(\ell)} - m^{(k)}\|_2 \le \|\chi_{-\ell+1}^{-k} - \chi_{-\ell}^{-k-1} \circ T\|_2 + \|\chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T\|_2. \tag{4.7}$$

We deal with these two terms separately. For the first term of (4.7), calculate that

$$\begin{aligned}\|\chi_{-\ell+1}^{-k} - \chi_{-\ell}^{-k-1} \circ T\|_2^2 &= \mathbb{E}\left(\left(\chi_{-\ell+1}^{-k} - \chi_{-\ell}^{-k-1} \circ T\right)^2\right)\\ &= \mathbb{E}\left(\left(\chi_{-\ell+1}^{-k}\right)^2 - 2\chi_{-\ell+1}^{-k}\,\chi_{-\ell}^{-k-1} \circ T + \left(\chi_{-\ell}^{-k-1}\right)^2\right). \end{aligned}\tag{4.8}$$

Considering the cross term in (4.8), we see

$$\mathbb{E}\left(\chi_{-\ell+1}^{-k}\,\chi_{-\ell}^{-k-1} \circ T\right) = \mathbb{E}\mathbb{E}_1\left(\chi_{-\ell+1}^{-k}\,\chi_{-\ell}^{-k-1} \circ T\right) = \mathbb{E}\left(\mathbb{E}_1\left(\chi_{-\ell+1}^{-k}\right)\chi_{-\ell}^{-k-1} \circ T\right).$$

Noting $\mathbb{E}_1\left(\chi_{-\ell+1}^{-k}\right) = \mathbb{E}_1\left(\sum_{j=-\ell+1}^{-k} \mathbb{E}_0 V_j\right) = \sum_{j=-\ell+1}^{-k}(\mathbb{E}_0 V_{j-1}) \circ T = \chi_{-\ell}^{-k-1} \circ T$, it follows that the cross term in (4.8) is

$$\mathbb{E}\left(\chi_{-\ell+1}^{-k}\,\chi_{-\ell}^{-k-1} \circ T\right) = \mathbb{E}\left(\chi_{-\ell}^{-k-1} \circ T\,\chi_{-\ell}^{-k-1} \circ T\right) = \mathbb{E}\left(\left(\chi_{-\ell}^{-k-1}\right)^2\right).$$

Therefore returning to (4.8) we obtain

$$\begin{aligned}\|\chi_{-\ell+1}^{-k} - \chi_{-\ell}^{-k-1} \circ T\|_2^2 &= \mathbb{E}\left(\left(\chi_{-\ell+1}^{-k}\right)^2 - \left(\chi_{-\ell}^{-k-1}\right)^2\right)\\ &= \mathbb{E}\left(\left(\chi_{-\ell+1}^{-k} - \chi_{-\ell}^{-k-1}\right)\left(\chi_{-\ell+1}^{-k} + \chi_{-\ell}^{-k-1}\right)\right)\\ &= \mathbb{E}\left(\left(a_{-k} - a_{-\ell}\right)\left(\chi_{-\ell+1}^{-k} + \chi_{-\ell}^{-k-1}\right)\right) \le 2\|v\|_\infty\|\chi_{-\ell+1}^{-k} + \chi_{-\ell}^{-k-1}\|_1\\ &\ll \|\chi_{-\ell+1}^{-k}\|_1 + \|\chi_{-\ell}^{-k-1}\|_1 \le 2\sum_{n=-\ell}^{-k}\|E_0 V_n\|_1 \to 0,\end{aligned}$$

as $k, \ell \to \infty$. It remains to show that the second term of (4.7) converges to zero. Note that

$$\begin{aligned}\|\chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T\|_2^2 &= \mathbb{E}\left(\left(\chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T\right)^2\right)\\ &= \mathbb{E}\left(\chi_{k+2}^{\ell+1}\left(\chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T\right)\right) - \mathbb{E}\left(\chi_{k+1}^{\ell} \circ T\left(\chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T\right)\right)\\ &= \mathbb{E}\left(\chi_{k+2}^{\ell+1}\left(\chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T\right)\right) + \mathbb{E}\left(\left(\chi_{k+1}^{\ell}\right)^2 - \chi_{k+1}^{\ell} \circ T\;\chi_{k+2}^{\ell+1}\right).\end{aligned}\tag{4.9}$$

We claim that the first term of (4.9) is zero. Since $\mathbb{E}_0\chi_{k+2}^{\ell+1} = 0$, the claim easily follows by showing $\chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T$ is $\mathcal{A}_0$-measurable. Indeed we have

$$\chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T = \sum_{j=k+1}^{\ell} \mathbb{E}_0 V_{j+1} - V_{j+1} - (\mathbb{E}_0 V_j) \circ T + V_{j+1} = \sum_{j} \mathbb{E}_0 V_{j+1} - (\mathbb{E}_0 V_j) \circ T,$$

which proves the claim. Next, for the second term of (4.9), add and subtract $\chi_{k+1}^{\ell}\ \chi_{k+2}^{\ell+1}$ to obtain

$$\begin{aligned}
\mathbb{E}\left(\left(\chi_{k+1}^{\ell}\right)^2 - \chi_{k+1}^{\ell} \circ T\ \chi_{k+2}^{\ell+1}\right) &= \mathbb{E}\left(\left(\chi_{k+1}^{\ell}\right)^2 - \chi_{k+1}^{\ell}\ \chi_{k+2}^{\ell+1} + \chi_{k+1}^{\ell}\ \chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T\ \chi_{k+2}^{\ell+1}\right) \\
&= \mathbb{E}\left(\chi_{k+1}^{\ell}(\chi_{k+1}^{\ell} - \chi_{k+2}^{\ell+1}) + (\chi_{k+1}^{\ell} - \chi_{k+1}^{\ell} \circ T)\chi_{k+2}^{\ell+1}\right) \\
&= \mathbb{E}\left(\chi_{k+1}^{\ell}(a_{k+1} - a_{\ell+1})\right) + \mathbb{E}\left((\chi_{k+1}^{\ell} - \chi_{k+1}^{\ell} \circ T)\chi_{k+2}^{\ell+1}\right).
\end{aligned}$$

Equation (4.9) now becomes

$$\begin{aligned}
\|\chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T\|_2^2 &= \mathbb{E}\left(\chi_{k+1}^{\ell}(a_{k+1} - a_{\ell+1})\right) + \mathbb{E}\left((\chi_{k+1}^{\ell} - \chi_{k+1}^{\ell} \circ T)\chi_{k+2}^{\ell+1}\right) \\
&\leq 4\|v\|_\infty \|\chi_{k+1}^{\ell}\|_1 + \mathbb{E}\left((\chi_{k+1}^{\ell} - \chi_{k+1}^{\ell} \circ T)\chi_{k+2}^{\ell+1}\right).
\end{aligned} \tag{4.10}$$

Next we focus on the second term of (4.10). Note that

$$\begin{aligned}
\chi_{k+1}^{\ell} \circ T - \chi_{k+1}^{\ell} &= \sum_{j=k+1}^{\ell} ((\mathbb{E}_0 V_j) \circ T - \mathbb{E}_0 V_j - V_{j+1} + V_j) \\
&= \sum_{j=k+1}^{\ell} ((\mathbb{E}_0 V_j) \circ T - \mathbb{E}_0 V_j) + V_{\ell+1} - V_{k+1} =: W + V_{\ell+1} - V_{k+1},
\end{aligned}$$

where W is $\mathcal{A}_0$-measurable. Since $\mathbb{E}_0 \chi_{k+2}^{\ell+1} = 0$ we see the second term of (4.10) is

$$\begin{aligned}
\mathbb{E}\left(\left(\chi_{k+1}^{\ell} \circ T - \chi_{k+1}^{\ell}\right)\chi_{k+2}^{\ell+1}\right) &= \mathbb{E}\mathbb{E}_0\left(W\ \chi_{k+2}^{\ell+1}\right) + \mathbb{E}\left(\left(V_{\ell+1} - V_{k+1}\right)\chi_{k+2}^{\ell+1}\right) \\
&= \mathbb{E}\left(\left(V_{\ell+1} - V_{k+1}\right)\chi_{k+2}^{\ell+1}\right).
\end{aligned}$$

Therefore (4.10) is now

$$\begin{aligned}
\|\chi_{k+2}^{\ell+1} - \chi_{k+1}^{\ell} \circ T\|_2^2 &\leq 4\|v\|_\infty \|\chi_{k+1}^{\ell}\|_1 + \mathbb{E}\left((V_{k+1} - V_{\ell+1})\chi_{k+2}^{\ell+1}\right) \\
&\leq 4\|v\|_\infty \left\|\chi_{k+1}^{\ell}\right\|_1 + 2\|v\|_\infty \left\|\chi_{k+2}^{\ell+1}\right\|_1 \ll \sum_{n=k+1}^{\ell+1} \|\mathbb{E}_0 V_n - V_n\|_1 \to 0,
\end{aligned}$$

as $k, \ell \to \infty$. Hence $m^{(k)}$ is a Cauchy sequence in $L^2(\Lambda)$ and so $m^{(k)} \to m$ in $L^2(\Lambda)$ and $m \in L^2(\Lambda)$. □

Proposition 4.1.4 (L^2 orthogonality of martingale differences)**.** *Suppose $f : \Lambda \to \mathbb{R}$ is an $L^2(\Lambda)$ observable which is $\mathcal{A}_0$-measurable and $\mathbb{E}_1 f = 0$. Then*

$$\int f \circ T^i \, f \circ T^j \, d\mu = \delta_{ij} \int f^2 \, d\mu$$

for every $i, j \in \mathbb{Z}$, where δ_{ij} is the Kronecker delta.

Proof. The case $i = j$ is clear from invariance of T. Without loss assume $i > j$ and note $f \circ T^{i-j}$ is $\mathcal{A}_1$-measurable. Now calculate

$$\int f \circ T^i \, f \circ T^j \, d\mu = \int f \, f \circ T^{i-j} \, d\mu = \int \mathbb{E}_1\big(f \, f \circ T^{i-j}\big) \, d\mu = \int \mathbb{E}_1 f \, f \circ T^{i-j} \, d\mu = 0,$$

and the result is shown. □

For square integrable, $\mathcal{A}_0$-measurable observables $f : \Lambda \to \mathbb{R}$ such that $\mathbb{E}_1 f = 0$, Proposition 4.1.4 shows

$$\left\| \sum_{j=0}^{n-1} f \circ T^j \right\|_2 = n^{1/2} \|f\|_2.$$

Preliminary Calculations

The invertible limit theorems follow the same proof structure as their non-invertible counterparts, with only a few calculations being different. We do these calculations now for ease of exposition.

Recall that $\mathcal{A}_0$ is a sub-sigma algebra of $\mathcal{A}$ that coarsens under T: With the notation $\mathcal{A}_n = T^{-n}\mathcal{A}_0$, $\mathbb{E}_n v = \mathbb{E}(v \,|\, \mathcal{A}_n)$ we have $\cdots \subseteq \mathcal{A}_1 \subseteq \mathcal{A}_0 \subseteq \mathcal{A}_{-1} \subseteq \cdots$ and for any $m, \ell \in \mathbb{Z}$,

$$(\mathbb{E}_m v) \circ T^\ell = \mathbb{E}_{m+\ell}(v \circ T^\ell), \quad \mathbb{E}_m \mathbb{E}_\ell v = \mathbb{E}_{\max\{m,\ell\}} v.$$

Proposition 4.1.5. *We have that*

$$\lim_{k \to \infty} \sup_{n \geq 1} n^{-1/2} \left\| \overline{\big(\mathbb{E}_0(v \circ T^{-k})\big)_n^-} \right\|_2 = 0.$$

Proof. Since $\mathbb{E}_0(v \circ T^{-1})$ plays the role that the transfer operator did in the non-invertible case, the proof used for the non-invertible analogue (Proposition 3.3.11) carries over with a few minor changes. Fix $n \geq 1$ and consider the decomposition $\mathbb{E}_0(v \circ T^{-k}) = r^{(k)} + \psi \circ T - \psi$, where

$$\psi = \sum_{\ell=1}^{\infty} \mathbb{E}_0\big(\mathbb{E}_0(v \circ T^{-k}) \circ T^{-\ell}\big) = \sum_{\ell}(\mathbb{E}_{\ell+k}\mathbb{E}_k v) \circ T^{-\ell-k} = \sum_{\ell}(\mathbb{E}_{\ell+k} v) \circ T^{-\ell-k}$$
$$= \sum_{\ell=1}^{\infty} \mathbb{E}_0(v \circ T^{-\ell-k}) = \sum_{\ell=k+1}^{\infty} \mathbb{E}_0(v \circ T^{-\ell}).$$

It follows that $r^{(k)}$ is $\mathcal{A}_0$-measurable and we see $\mathbb{E}_1 r^{(k)} = 0$ since

$$\begin{aligned}\mathbb{E}_1 r^{(k)} &= \mathbb{E}_1\big(\mathbb{E}_0(v \circ T^{-k}) + \psi - \psi \circ T\big) \\ &= \mathbb{E}_1(v \circ T^{-k}) + \mathbb{E}_1\left(\sum_{\ell=k+1}^{\infty} \mathbb{E}_0(v \circ T^{-\ell}) - \mathbb{E}_0(v \circ T^{-\ell}) \circ T\right) \\ &= \mathbb{E}_1(v \circ T^{-k}) + \sum_{\ell=k+1}^{\infty} \mathbb{E}_1(v \circ T^{-\ell}) - \mathbb{E}_1(v \circ T^{-\ell+1}) = 0.\end{aligned}$$

As $\mathbb{E}_0(v \circ T^{-k}) \circ T^{-j}$ is $\mathcal{A}_{-j}$-measurable, we apply Rio's inequality (Proposition 3.3.10) with the filtration $\mathcal{A}_{-j},\ j \geq 1$. Denoting

$$b_{j,\ell} = \left\| \mathbb{E}_0(v \circ T^{-k}) \circ T^{-j} \sum_{u=j}^{\ell} \mathbb{E}\big(\mathbb{E}_0(v \circ T^{-k}) \circ T^{-u} \big| \mathcal{A}_{-j}\big) \right\|_1,$$

we have

$$\left\| \overline{\big(\mathbb{E}_0(v \circ T^{-k})\big)_n^-} \right\|_2^2 = \mathbb{E}\left(\overline{\big(\mathbb{E}_0(v \circ T^{-k})\big)_n^-}\right)^2 \leq C \sum_{j=1}^{n} \max_{1 \leq \ell \leq n} b_{j,\ell}. \tag{4.11}$$

Note that

$$\begin{aligned} b_{j,\ell} &\leq \|v\|_\infty \left\| \sum_{u=j}^{\ell} \mathbb{E}_{-j}\Big((r^{(k)} + \psi \circ T - \psi) \circ T^{-u} \Big) \right\|_1 \\ &\leq \|v\|_\infty \left\| \sum_{u=j}^{\ell} \big(\mathbb{E}_{-j+u} r^{(k)}\big) \circ T^{-u} \right\|_1 + \|v\|_\infty \left\| \mathbb{E}_{-j}\left(\sum_{u=j}^{\ell} (\psi \circ T - \psi) \circ T^{-u} \right) \right\|_1 \\ &\leq \|v\|_\infty \sum_{u=j}^{\ell} \big\|\mathbb{E}_{-j+u} r^{(k)}\big\|_1 + \|v\|_\infty \big\|\psi \circ T^{-j+1} - \psi \circ T^{-\ell}\big\|_1. \end{aligned}$$

For any $p \geq 1$ we have $\mathbb{E}_p r^{(k)} = \mathbb{E}_p \mathbb{E}_1 r^{(k)} = 0$, and so

$\sum_{u=j}^{\ell} \left\|\mathbb{E}_{-j+u} r^{(k)}\right\|_1 = \|\mathbb{E}_0 r^{(k)}\|_1 = \|r^{(k)}\|_1$. Therefore

$$b_{j,\ell} \leq \|v\|_\infty \big(\|r^{(k)}\|_1 + 2\|\psi\|_1\big).$$

Returning to (4.11) we now have

$$\begin{aligned}
\left\|\overline{\left(\mathbb{E}_0(v \circ T^{-k})\right)_n^-}\right\|_2^2 &\ll \sum_{j=1}^{n} \max_{1 \leq \ell \leq n} \Big(\|r^{(k)}\|_1 + 2\|\psi\|_1\Big) \\
&= n\big(\|E_0(v \circ T^{-k}) - \psi \circ T + \psi\|_1 + 2\|\psi\|_1\big) \\
&\leq n\left(\|E_0(v \circ T^{-k})\|_1 + 4\left\|\sum_{\ell=k+1}^{\infty} \mathbb{E}_0(v \circ T^{-\ell})\right\|_1\right) \\
&\ll n \sum_{\ell=k}^{\infty} \|\mathbb{E}_0(v \circ T^{-\ell})\|_1.
\end{aligned}$$

Hence

$$\sup_n n^{-1/2} \left\|\overline{\left(\mathbb{E}_0(v \circ T^{-k})\right)_n^-}\right\|_2 \ll \left(\sum_{\ell=k}^{\infty} \|\mathbb{E}_0(v \circ T^{-\ell})\|_1\right)^{1/2} \to 0,$$

as $k \to \infty$. □

The following proposition does not require the random variables to be adapted to any filtration.

Proposition 4.1.6 ([11, Equation (3.4)]). *For a sum $S_n = \sum_{j=1}^{n} X_j$ of square integrable random variables we have*

$$\mathbb{E}\big((S_n^*)^2\big) \leq 4\mathbb{E}(S_n^2) - 4\sum_{j=1}^{n} \mathbb{E}(X_j\, S_{j-1}^*).$$

□

Proposition 4.1.7. *We have that*

$$\lim_{k\to\infty} \sup_n n^{-1/2} \left\|\overline{\left(\mathbb{E}_0(v \circ T^k) - v \circ T^k\right)_n^-}\right\|_2 = 0.$$

The proof of Proposition 4.1.7 is similar to an argument presented in [11, Proposition 4].

Proof. Let $k \geq 1$, $V_j := v \circ T^j$, $X_j := \mathbb{E}_{-k-j} V_{-j} - V_{-j}$ and $S_n := \sum_{j=1}^{n} X_j$. Without loss assume that $n > k+1$. Define $S_n^* := \max\{0, S_1, \ldots, S_n\}$, $S_{n,*} :=$

$\max\{0, -S_1, \ldots, -S_n\}$ and recall that $\overline{S_n} = \max_{0\le \ell \le n} |S_\ell|$. We want to show

$$\lim_{n\to\infty} n^{-1}\mathbb{E}\left(\overline{S_n}^2\right) = 0.$$

Since $\overline{S_n}^2 = \big(\max\{S_n^*, S_{n,*}\}\big)^2 = \max\{\big(S_n^*\big)^2, \big(S_{n,*}\big)^2\} \le (S_n^*)^2 + (S_{n,*})^2$, it suffices to show

$$\lim_{n\to\infty} n^{-1}\left[\mathbb{E}\big((S_n^*)^2\big) + \mathbb{E}\big((S_{n,*})^2\big)\right] = 0.$$

We first bound $n^{-1}\mathbb{E}\big((S_n^*)^2\big)$. By Proposition 4.1.6,

$$\mathbb{E}\big((S_n^*)^2\big) \le 4\mathbb{E}(S_n^2) - 4\sum_{j=1}^{n} \mathbb{E}(X_j\, S_{j-1}^*) \le 4\mathbb{E}(S_n^2) + 4\sum_{j=1}^{n} \left|\mathbb{E}(X_j\, S_{j-1}^*)\right|. \tag{4.12}$$

Note that $X_j = (E_{-k}v - v)\circ T^{-j}$. For the first term of (4.12), the calculations from the proof of Proposition 3.2.7 show

$$\begin{aligned}\mathbb{E}(S_n^2) &= \sum_{j=1}^{n} \int X_j^2\, d\mu + 2\sum_{r=1}^{n}(n-r)\int X_0\, X_{-r}\, d\mu \\ &\le 2n\|v\|_\infty \|\mathbb{E}_{-k}v - v\|_1 + 2n\sum_{r}\left|\int (\mathbb{E}_{-k}v - v)\ (\mathbb{E}_{-k}v - v)\circ T^r\, d\mu\right|. \end{aligned}\tag{4.13}$$

Next, note that $(\mathbb{E}_{-k}v)\circ T^r$ is $\mathcal{A}_{-k+r}$-measurable and hence $\mathcal{A}_{-k}$-measurable. Similarly $(\mathbb{E}_{-k-r}v)\circ T^r$ is $\mathcal{A}_{-k}$-measurable. Therefore

$$\begin{aligned}&\int \big(\mathbb{E}_{-k}v - v\big)\ \big(\mathbb{E}_{-k-r}v - \mathbb{E}_{-k}v\big)\circ T^r\, d\mu = \int \mathbb{E}_{-k}\Big[\big(\mathbb{E}_{-k}v - v\big)\ \big(\mathbb{E}_{-k-r}v - \mathbb{E}_{-k}v\big)\circ T^r\Big]\, d\mu \\ &= \int \mathbb{E}_{-k}\big(\mathbb{E}_{-k}v - v\big)\ \big(\mathbb{E}_{-k-r}v - \mathbb{E}_{-k}v\big)\circ T^r\, d\mu = 0.\end{aligned}$$

In particular, $\int \big(\mathbb{E}_{-k}v - v\big)\big(\mathbb{E}_{-k}v\big)\circ T^r\, d\mu = \int \big(\mathbb{E}_{-k}v - v\big)\big(\mathbb{E}_{-k-r}v\big)\circ T^r\, d\mu$ and, returning to (4.13), we have

$$\begin{aligned}\mathbb{E}(S_n^2) &\ll n\|\mathbb{E}_{-k}v - v\|_1 + n\sum_{r=1}^{n}\left|\int (\mathbb{E}_{-k}v - v)\ (\mathbb{E}_{-k-r}v - v)\circ T^r\, d\mu\right| \\ &\le n\|\mathbb{E}_{-k}v - v\|_1 + 2\|v\|_\infty n\sum_{r=1}^{n}\|\mathbb{E}_{-k-r}v - v\|_1 \ll n\sum_{r=k}^{\infty}\|\mathbb{E}_{-r}v - v\|_1.\end{aligned}\tag{4.14}$$

Next, we bound the second term of (4.12). Define $Y_{i,j} := \mathbb{E}_{-k-i}V_{-i} - \mathbb{E}_{-k-j}V_{-i}$ and $T_{p,j} := \sum_{i=1}^{p} Y_{i,j}$. For every $i < j$ we have $\mathcal{A}_{-k-i} \subseteq \mathcal{A}_{-k-j}$ and so $Y_{i,j}$ is $\mathcal{A}_{-k-j}$-measurable. Therefore $T_{j-1}^* := \max\{0, T_{1,j}, \ldots, T_{j-1,j}\}$ is $\mathcal{A}_{-k-j}$-

measurable. Note that $\mathbb{E}(X_j T^*_{j-1}) = \mathbb{E}\mathbb{E}_{-k-j}(X_j T^*_{j-1}) = \mathbb{E}\big(\mathbb{E}_{-k-j}\big(\mathbb{E}_{-k-j}V_{-j} - V_{-j}\big)T^*_{j-1}\big) = 0$. Hence

$$\sum_{j=1}^{n} \big|\mathbb{E}(X_j S^*_{j-1})\big| = \sum_j \Big|\mathbb{E}\Big(X_j(S^*_{j-1} - T^*_{j-1})\Big)\Big| \le 2\|v\|_\infty \sum_j \mathbb{E}\big|S^*_{j-1} - T^*_{j-1}\big|.$$

Now, by Proposition 2.2.5 we have

$$|S^*_{j-1} - T^*_{j-1}| \le \sum_{i=1}^{j-1} |X_i - Y_{i,j}| = \sum_i |\mathbb{E}_{-k-j}V_{-i} - V_{-i}|.$$

Therefore we obtain

$$\begin{aligned}\sum_{j=1}^{n} \big|\mathbb{E}(X_j\, S^*_{j-1})\big| &\le 2\|v\|_\infty \sum_{j=1}^{n}\sum_{i=1}^{j-1} \mathbb{E}\big|\mathbb{E}_{-k-j}V_{-i} - V_{-i}\big| = 2\|v\|_\infty \sum_{1\le i<j\le n} \|\mathbb{E}_{-k-j+i}v - v\|_1 \\ &= 2\|v\|_\infty \sum_{r=1}^{n}(n-r)\|\mathbb{E}_{-k-r}v - v\|_1 \qquad (4.15)\end{aligned}$$

Combining (4.12) with (4.14) and (4.15) we see that

$$\begin{aligned}n^{-1}\mathbb{E}\big((S^*_n)^2\big) &\ll n^{-1}\left(n\sum_{r=k}^{\infty}\|\mathbb{E}_{-r}v - v\|_1 + 2\|v\|_\infty \sum_{r=1}^{n}(n-r)\|\mathbb{E}_{-k-r}v - v\|_1\right) \\ &\ll \sum_{r=k}^{\infty}\|\mathbb{E}_{-r}v - v\|_1.\end{aligned}$$

To obtain a bound for $(S_{n,*})^2$, we note that changing X_j to $-X_j$ changes S_n to $-S_n$ and S^*_n to $S_{n,*}$. Therefore, the corresponding bound for $(S_{n,*})^2$ is also

$$n^{-1}\mathbb{E}\big((S_{n,*})^2\big) \ll \sum_{r=k}^{\infty}\|\mathbb{E}_{-r}v - v\|_1.$$

The result now follows since

$$\sup_{n\ge 1}\big[n^{-1}\mathbb{E}\big((S^*_n)^2\big) + n^{-1}\mathbb{E}\big((S_{n,*})^2\big)\big] \ll \sup_n \sum_{r=k}^{\infty}\|\mathbb{E}_{-r}v - v\|_1 \to 0,$$

as $k \to \infty$. $\square$

Proposition 4.1.8. *We have that*

$$\lim_{n\to\infty} n^{-1/2}\left\|\overline{(v-m)}^{-}_{n}\right\|_2 = 0.$$

Proof. First note that

$$n^{-1/2}\left\|\overline{(v-m)_n^-}\right\|_2 \le n^{-1/2}\left\|\overline{\left(m-m^{(k)}\right)_n^-}\right\|_2 + n^{-1/2}\left\|\overline{\left(v-m^{(k)}\right)_n^-}\right\|_2. \qquad (4.16)$$

Note $(m-m^{(k)})\circ T^{-j}$ forms a sequence of martingale differences by Proposition 4.2.1. For the first term of (4.16), use Doob's inequality (Proposition 3.3.9) and the orthogonality of martingale differences (Proposition 4.1.4) to see

$$n^{-1/2}\left\|\overline{\left(m-m^{(k)}\right)_n^-}\right\|_2 \le 2n^{-1/2}\left\|\left(m-m^{(k)}\right)_n^-\right\|_2 = 2\|m-m^{(k)}\|_2.$$

Proposition 4.1.3 shows $\|m-m^{(k)}\|_2 \to 0$ as $k\to\infty$ and so it remains to bound the second term of (4.16). To do this, we use the truncated decomposition $v = m^{(k)} + \chi_{-k}^k \circ T - \chi_{-k}^k + \mathbb{E}_0(v\circ T^{-k}) - \mathbb{E}_0(v\circ T^{k+1}) + v\circ T^{k+1}$ and take $\limsup_{n\to\infty}$ to see

$$\begin{aligned}\limsup_{n\to\infty} n^{-1/2}\left\|\overline{\left(v-m^{(k)}\right)_n^-}\right\|_2 &\le \limsup_{n\to\infty} n^{-1/2}\left\|\overline{\left(\mathbb{E}_0(v\circ T^{k+1}) - v\circ T^{k+1}\right)_n^-}\right\|_2 \\ &\quad + \limsup_{n\to\infty} n^{-1/2}\left\|\overline{\left(\mathbb{E}_0(v\circ T^{-k})\right)_n^-}\right\|_2 \\ &\quad + \limsup_{n\to\infty} n^{-1/2}\left\|\overline{\left(\chi_{-k}^k\circ T - \chi_{-k}^k\right)_n^-}\right\|_2.\end{aligned}$$

The first two terms converge to zero as $k\to\infty$ by Propositions 4.1.5 and 4.1.7. It remains to show

$$\lim_{k\to\infty}\limsup_{n\to\infty} n^{-1/2}\left\|\overline{\left(\chi_{-k}^k\circ T - \chi_{-k}^k\right)_n^-}\right\|_2 = 0.$$

For any $k\ge 1$ we have $|\chi_{-k}^k| \le (3k+2)\|v\|_\infty$, and so

$$\limsup_{n\to\infty} n^{-1/2}\left\|\overline{\left(\chi_{-k}^k\circ T - \chi_{-k}^k\right)_n^-}\right\|_2 = \limsup_n n^{-1/2}\left\|\max_{1\le \ell\le n}\left|\sum_{j=1}^{\ell}\left(\chi_{-k}^k\circ T - \chi_{-k}^k\right)\circ T^{-j}\right|\right\|_2$$
$$= \limsup_n n^{-1/2}\left\|\max_\ell\left|\chi_{-k}^k - \chi_{-k}^k\circ T^{-\ell}\right|\right\|_2 \le \limsup_n n^{-1/2}(6k+4)\|v\|_\infty = 0.$$

□

4.2 Central Limit Theorem

Proposition 4.2.1. *Suppose $f\in L^1(\Lambda)$ is $\mathcal{A}_0$-measurable and $\mathbb{E}_1 f = 0$. Then $f_n^- = \sum_{j=1}^n f\circ T^{-j}$ is a martingale with respect to the filtration $\mathcal{A}_{-n} = T^n\mathcal{A}_0$.*

Proof. Firstly we note that $\|f_n^-\|_1 \le n\|f\|_1 < \infty$. Secondly, f_n^- is $\mathcal{A}_{-n}$-measurable since $f \circ T^{-j}$ is $\mathcal{A}_{-j}$-measurable and $\mathcal{A}_{-j} \subseteq \mathcal{A}_{-n}$ for all $j \le n$. Finally, we have

$$\mathbb{E}(f \circ T^{-j} \,|\, \mathcal{A}_{-j+1}) = \mathbb{E}_{-j+1}(f \circ T^{-j}) = \mathbb{E}_1(f) \circ T^{-j} = 0,$$

so $f \circ T^{-j}$ forms a martingale difference sequence and f_n^- is a martingale with respect to $\mathcal{A}_{-n}$. $\square$

Proposition 4.2.2 (Variance)**.** *We have that*

$$\lim_{n\to\infty} \frac{1}{n} \int v_n^2 \, d\mu = \int m^2 \, d\mu.$$

Proof. By orthogonality of martingale differences (Proposition 4.1.4) and Proposition 4.1.8 we have

$$\begin{aligned} \left| n^{-1/2}\|v_n\|_2 - \|m\|_2 \right| &= \left| n^{-1/2}\|v_n\|_2 - n^{-1/2}\|m_n\|_2 \right| \le n^{-1/2}\|(v-m)_n\|_2 \\ &= n^{-1/2}\|(v-m)_n \circ T^{-n}\|_2 = n^{-1/2}\|(v-m)_n^-\|_2 \\ &\le n^{-1/2}\|\overline{(v-m)_n^-}\|_2 \to 0. \end{aligned}$$

$\square$

Theorem 4.2.3 (Scalar CLT)**.** *We have*

$$n^{-1/2} v_n \to_d \mathcal{N}(0, \sigma^2),$$

where $\sigma^2 := \lim_{n\to\infty} n^{-1} \int v_n^2 \, d\mu$. Moreover, $\sigma^2 = 0$ if and only if v is an $L^1(\Lambda)$ coboundary.

Remark 1. This result is not new; Instead of Assumption ($\mathrm{A_{inv}}$), Liverani [32, Theorem 1.2] assumes the conditions

(i) $\displaystyle\sum_{n=0}^{\infty} \left| \int v \, v \circ T^n \, d\mu \right| < \infty,$

(ii) $\displaystyle\sum_{n=0}^{\infty} \left| \mathbb{E}_0(v \circ T^{-n}) \right|$ converges in $L^1(\Lambda)$,

(iii) There exists $\alpha > 1$ such that $\displaystyle\sup_{k\ge 1} k^\alpha \|\mathbb{E}_0(v \circ T^n) - v \circ T^n\|_1 < \infty.$

Remark 2. Zweimüller [54, Corollary 1] proved that the CLT also holds with respect to every probability measure ν which is absolutely continuous with respect to μ.

Proof of Theorem 4.2.3. We see m_n^- is a martingale due to Proposition 4.2.1. Proposition 3.2.6 now shows $n^{-1/2}m_n^- \to_d \mathcal{N}(0,\sigma^2)$ with $\sigma^2 = \int m^2\, d\mu$. Proposition 4.2.2 shows $\sigma^2 = \lim_n n^{-1}\int v_n^2\, d\mu$. The CLT for $n^{-1/2}v_n$ now follows since $m_n^- = m_n \circ T^{-n} =_d m_n$ and $n^{-1/2}\|v_n - m_n\|_1 = n^{-1/2}\|\chi \circ T^n - \chi\|_1 \le 2n^{-1/2}\|\chi\|_1 \to 0$ as $n \to \infty$. We easily see $\sigma^2 = \|m\|_2^2 = 0$ if and only if $m = 0$, so v is an $L^1(\Lambda)$ coboundary. $\square$

Example 4.2.4. Suppose $0 < \alpha < 1/2$, $\Lambda = [0,1] \times [0,1]$ and let $T_{\alpha\alpha} : \Lambda \to \Lambda$ be the intermittent Baker's map defined in (4.3). Let $v : \Lambda \to \mathbb{R}$ be a Hölder continuous and mean zero observable. Proposition 4.0.5 shows

$$\|\mathbb{E}_0(v \circ T_{\alpha\alpha}^{-n})\|_1 \le Cn^{-1/\alpha+1}, \quad \|\mathbb{E}_0(v \circ T_{\alpha\alpha}^{n}) - v \circ T_{\alpha\alpha}^{n}\|_1 \le Cn^{-1/\alpha+1},$$

and so assumption ($\mathrm{A_{inv}}$) is satisfied for all $0 < \alpha < 1/2$. The CLT now follows from Theorem 4.2.3.

As in the non-invertible case, Assumption ($\mathrm{A_{inv}}$) and the decompositions in Proposition 4.1.1 and 4.1.2 easily carry over to vector-valued observables $v \in L^\infty(\Lambda, \mathbb{R}^d)$ by considering them component wise.

Proposition 4.2.5 (Covariance)**.** *For $v \in L^\infty(\Lambda, \mathbb{R}^d)$ we have that*

$$\lim_{n\to\infty} \frac{1}{n}\int v_n\, v_n^T\, d\mu = \int m\, m^T\, d\mu.$$

Proof. The proof for non-invertible maps (Proposition 3.2.14) carries over verbatim after replacing the non-invertible scalar variance (Proposition 3.2.8) with its invertible counterpart (Proposition 4.2.2). $\square$

Theorem 4.2.6 (Vector CLT)**.** *We have*

$$n^{-1/2}v_n \to_d \mathcal{N}(0,\Sigma),$$

where $\Sigma := \lim_{n\to\infty} n^{-1}\int v_n v_n^T\, d\mu$. For the degenerate case, $\det\Sigma = 0$ if and only if there exists a non-zero $c \in \mathbb{R}^d$ such that $c^T v = \chi \circ T - \chi$ for some $\chi \in L^1(\Lambda)$.

Proof. The covariance $\Sigma = \int m\, m^T\, d\mu = \lim_{n\to\infty} n^{-1}\int v_n\, v_n^T\, d\mu$ is now due to Proposition 4.2.5. For any $c \in \mathbb{R}^d$, we see $c^T v \in L^\infty(\Lambda)$, $\int c^T v\, d\mu = 0$ and $\sum_{j\ge0} \big\|\mathbb{E}_0\big((c^T v) \circ T^{-j-1}\big)\big\|_1 + \big\|\mathbb{E}_0\big((c^T v) \circ T^j\big) - (c^T v) \circ T^j\big\|_1 < \infty$. Therefore the scalar CLT (Theorem 4.2.3) holds for $n^{-1/2}(c^T v)_n$ and the proof of the non-invertible vector CLT (Theorem 3.2.12) carries over. $\square$

4.3 Weak Invariance Principle

Define the forward càdlàg process

$$W_n(t) := n^{-1/2} \sum_{j=0}^{[nt]-1} v \circ T^j,$$

and the backwards càdlàg processes

$$W_n^-(t) := n^{-1/2} \sum_{j=1}^{[nt]} v \circ T^{-j}, \quad M_n^-(t) := n^{-1/2} \sum_{j=1}^{[nt]} m \circ T^{-j}.$$

Proposition 4.3.1. *For each $n \geq 1$, $M_n^- : [0,1] \to \mathbb{R}^d$ is a martingale with respect to the filtration $\mathcal{A}_{n,t} = \{T^{[nt]}\mathcal{A}_0 \,|\, 0 \leq t \leq 1\}$.*

Proof. Let $n \geq 1$. Firstly, $M_n^-(t) \in L^1(\Lambda, \mathbb{R}^d)$ since

$$\|M_n^-(t)\|_1 = n^{-1/2} \left\| \sum_{j=1}^{[nt]} m \circ T^{-j} \right\|_1 \leq \frac{[nt]}{n^{1/2}} \|m\|_1 < \infty.$$

Secondly, $M_n^-(t)$ is $\mathcal{A}_{n,t}$-measurable since $m \circ T^{-j}$ is $\mathcal{A}_{-j}$-measurable and $\mathcal{A}_{-j} \subseteq \mathcal{A}_{n,t}$ for all $j \leq [nt]$. It remains to show that

$$\mathbb{E}(M_n^-(t) \,|\, \mathcal{A}_{n,s}) = M_n^-(s),$$

for all $s \leq t$. If $[nt] = [ns]$, then we have $M_n^-(t) = M_n^-(s)$ and we are finished. Suppose $[nt] > [ns]$ and note that

$$\mathbb{E}(n^{-1/2} \sum_{j=1}^{[nt]} m \circ T^{-j} \,|\, \mathcal{A}_{n,s}) = M_n^-(s) + n^{-1/2} \sum_{j=[ns]+1}^{[nt]} \mathbb{E}(m \circ T^{-j} \,|\, T^{[ns]}\mathcal{A}_0).$$

Let $j = [ns]+1, \ldots, [nt]$ be fixed, and calculate that

$$\mathbb{E}(m \circ T^{-j} \,|\, T^{[ns]}\mathcal{A}_0) = \mathbb{E}(m \,|\, T^{[ns]-j}\mathcal{A}_0) \circ T^{-j} = \big(\mathbb{E}_{j-[ns]}\mathbb{E}_1 m\big) \circ T^{-j} = 0.$$

Therefore $\mathbb{E}(M_n^-(t) \,|\, \mathcal{A}_{n,s}) = M_n^-(s)$ and M_n^- is a martingale. □

Theorem 4.3.2 (Invertible WIP)**.** *We have*

$$W_n \to_w W,$$

in $\big(D([0,1],\mathbb{R}^d),\|\cdot\|_\infty\big)$*, where* W *is a* d *dimensional Brownian Motion with mean zero and covariance* $\Sigma = \lim_{n\to\infty} n^{-1}\int v_n v_n^T\,d\mu$*.*

Remark 1. This was first proven by Dedecker and Rio [11, Corollary 4(b)].

Remark 2. As in Theorem 4.2.6, $\det\Sigma = 0$ if and only if there exists a non-zero $c\in\mathbb{R}^d$ such that $c^T v$ is an $L^1(\Lambda)$ coboundary.

Remark 3. We could equally define our processes in $D([0,K],\mathbb{R}^d)$ for any integer $K\geq 1$, and the result still holds. Hence the result is true if we define our processes in $D([0,\infty),\mathbb{R}^d)$ by [7, Section 16].

Remark 4. Zweimüller [54, Corollary 3] proved that the WIP also holds with respect to every probability measure ν which is absolutely continuous with respect to μ.

Proof of Theorem 4.3.2. By Proposition 4.3.1, M_n^- is a martingale. Since T is invertible, lifting is not needed and Proposition 3.3.8 carries over to the invertible case by simply dropping the tilde notation. Therefore we have $M_n^- \to_w W$ in $(D([0,1],\mathbb{R}^d),\|\cdot\|_\infty)$, where W is a d dimensional Brownian motion with mean zero and covariance $\Sigma = \int m\, m^T\,d\mu$. Proposition 4.2.5 shows $\Sigma = \lim_n n^{-1}\int v_n\, v_n^T\,d\mu$. Proposition 4.1.8 extends to vector-valued observables since for $f=(f^1,\dots,f^d)$ we have

$$\|f\|_2 = \big\|\|f\|_{\ell^1}\big\|_{L^2(\Lambda,\mathbb{R})} = \left\|\sum_{i=1}^d |f^i|\right\|_{L^2(\Lambda,\mathbb{R})} \leq \sum_i \|f^i\|_2.$$

Therefore $\lim_{n\to\infty} n^{-1/2}\left\|\overline{(v-m)}_n^-\right\|_2 = 0$ and

$$\sup_{t\in[0,1]}|W_n^-(t)-M_n^-(t)| = n^{-1/2}\max_{1\leq\ell\leq n}\left|\sum_{j=1}^{\ell}(v-m)\circ T^{-j}\right| = n^{-1/2}\ \overline{(v-m)_n^-} \to_p 0.$$

This gives the convergence $W_n^- \to_w W$. Now, the proof of the non-invertible WIP (Theorem 3.3.3) carries over by dropping the tilde notation and we see $W_n \to_w W$ as required. □

Example 4.3.3. Suppose $0<\alpha<1/2$, $\Lambda=[0,1]\times[0,1]$ and let $T_{\alpha\alpha}:\Lambda\to\Lambda$ be the intermittent Baker's map defined in (4.3). Let $v:\Lambda\to\mathbb{R}$ be a Hölder continuous and mean zero observable. Proposition 4.0.5 shows

$$\|\mathbb{E}_0(v\circ T_{\alpha\alpha}^{-n})\|_1 \leq Cn^{-1/\alpha+1}, \quad \|\mathbb{E}_0(v\circ T_{\alpha\alpha}^{n}) - v\circ T_{\alpha\alpha}^n\|_1 \leq Cn^{-1/\alpha+1},$$

and so assumption ($\mathrm{A_{inv}}$) is satisfied for all $0<\alpha<1/2$. The WIP now follows from Theorem 4.3.2.

4.4 Iterated Weak Invariance Principle

Recall the standing assumptions from the start of the chapter: Suppose $T : \Lambda \to \Lambda$ is an invertible, ergodic and measure-preserving transformation over the probability space $(\Lambda, \mathcal{A}, \mu)$ and $v \in L^\infty(\Lambda, \mathbb{R}^d)$ is a mean zero observable. In this section, we replace Assumption ($\mathrm{A_{inv}}$) with the following stronger one:

$$\sum_{n=0}^{\infty} \big\|\mathbb{E}_0(v \circ T^n) - v \circ T^n\big\|_2 < \infty, \qquad \sum_{n=1}^{\infty} \big\|\mathbb{E}_0(v \circ T^{-n})\big\|_1 < \infty. \tag{$\mathrm{A_{strong}}$}$$

Theorem 4.4.1 (Invertible Iterated WIP)**.** *Assume that $T : \Lambda \to \Lambda$ is also mixing. We have that*

$$\big(W_n, \mathbb{W}_n\big) \to_w \big(W, \mathbb{W}\big)$$

in $\big(D([0,1], \mathbb{R}^d \times \mathbb{R}^{d\times d}), \|\cdot\|_\infty\big)$, where W is a Brownian motion with mean zero and covariance $\Sigma = \lim_{n\to\infty} n^{-1} \int v_n v_n^T \, d\mu$, and

$$\mathbb{W}(t) = \int_0^t W \otimes dW + t \sum_{j=1}^{\infty} \int v\big(v \circ T^j\big)^T \, d\mu,$$

for $t \in [0,1]$.

Remark 1. This iterated WIP is already known by Kelly and Melbourne [29, Theorem 5.2] under the more restrictive assumption

$$\sum_{n=0}^{\infty} \big\|\mathbb{E}_0(v \circ T^n) - v \circ T^n\big\|_2 < \infty, \qquad \sum_{n=1}^{\infty} \big\|\mathbb{E}_0(v \circ T^{-n})\big\|_2 < \infty.$$

Remark 2. If T is modelled by a Young tower with polynomial tails $\mu(r > n) \leq Cn^{-\beta}$ for some $\beta > 2$ (and so assumption (A) holds by Theorem 4.0.4), then the iterated WIP is already known by Melbourne and Varandas [40, Corollary 2.3].

Remark 3. As in Theorem 3.2.12, $\det \Sigma = 0$ if and only if there exists a non-zero $c \in \mathbb{R}^d$ such that $c^T v$ is an $L^1(\Lambda)$ coboundary.

Remark 4. We could equally define our processes in $D([0,K], \mathbb{R}^d)$ for any integer $K \geq 1$, and the result still holds. Hence the result is true if we define our processes in $D([0,\infty), \mathbb{R}^d)$ by [7, Section 16].

Remark 5. It is still not known if the iterated WIP holds under the weaker assumption

$$sum_{n=0}^{\infty} \big\|\mathbb{E}_0(v \circ T^n) - v \circ T^n\big\|_1 < \infty, \qquad \sum_{n=1}^{\infty} \big\|\mathbb{E}_0(v \circ T^{-n})\big\|_1 < \infty.$$

Next we give an example where the iterated WIP holds, but is not covered by [29, 40]. Similar to the previous chapter, we first outline a generic situation where the mixing conditions for a suitable suspension is inherited from the underlying map which is modelled by a Young tower.

4.4.1 Suspensions and Examples

Suppose $T : \Lambda \to \Lambda$ is an invertible, ergodic and measure-preserving map over the probability space $(\Lambda, \mathcal{A}, \mu)$, and that T has a stable foliation $\mathcal{W}_T^s = \{W_T^s(x) \,:\, x \in \Lambda\}$. Define the suspension flow $T_t : \Lambda^h \to \Lambda^h$ as in (3.15) over an η-Hölder roof function h such that $\inf h > 0$. Further assume that the roof function h is constant on local stable leaves for T.

Proposition 4.4.2. *For any $t > 0$, the map T_t admits a partition into stable leaves given by $W^s(x, u) := W_T^s(x) \times \{u\}$.*

Proof. For any $(x, u) \in \Lambda^h$, define $W^s(x, u) := W_T^s(x) \times \{u\}$. Clearly this forms a partition of Λ^h. Let $N = N(x, u, t)$ be the lap number, i.e. the integer N satisfying $u+t \in [h_N(x), h_{N+1}(x))$. Then $T_t(x, u) = (T^N x, u+t-h_N(x))$. Since h is constant on stable leaves, $h(T^j x') = h(T^j x)$ and $h_n(x') = h_n(x)$ for every $x' \in W_T^s(x)$ and $n \geq 1$. Therefore $N(x', u, t) = N(x, u, t)$. Suppose that $(x', u) \in W^s(x, u) = W_T^s(x) \times \{u\}$ and $t > 0$. This partition is invariant under T_t since

$$\begin{aligned} T_t(x', u) = (T^N x', u+t-h_N(x')) &\in W_T^s(T^N x) \times \{u+t-h_N(x)\} \\ &= W^s((T^N x, u+t-h_N(x))) \\ &= W^s(T_t(x, u)). \end{aligned}$$

Moreover, we have

$$\begin{aligned} d(T_t(x', u), T_t(x, u)) &= d(T^{N(x',u,t)} x', T^{N(x,u,t)} x) + \big|u+t-h_N(x') - (u+t-h_N(x))\big| \\ &= d(T^{N(x,u,t)} x', T^{N(x,u,t)} x) \to 0, \end{aligned}$$

since $x' \in W_T^s(x)$. $\square$

Verifying Assumption ($\mathbf{A_{strong}}$)

Proposition 4.4.3 ([12])**.** *Let $v : \Lambda^h \to \mathbb{R}$ be a Hölder continuous, mean zero observable. Suppose T satisfies the hypothesis of Proposition 4.0.2, and assume*

$\inf_Y h > 1$ *and* $h \in L^\infty(\Lambda)$*. Then*

$$\|\mathbb{E}_0(v \circ T_1^n) - v \circ T_1^n\|_1 \leq Cn^{-\alpha+1}.$$

□

Remark. Roughly speaking, we define our inducing set for T_1 as $Y^h = \{(y,u) : y \in Y, u \in [0, h(y)]\}$. The assumption $\inf_Y h > 1$ ensures that the first return τ^h of T_1 to Y^h has tails comparable to the first return τ of T to Y. Without this assumption, h can be very small on Y and the time one map may 'pass over' the inducing set Y^h, meaning we cannot control the tail of τ^h. We can drop the assumption $\inf h > 1$ if we replace the time-one map T_1 with the time-($\inf h - \epsilon$) map $T_{\inf h - \epsilon}$, for any $\epsilon > 0$.

Proof. We verify the hypothesis of Proposition 4.0.2 for T_1. Define $Y^h := \{(y,u) : y \in Y, u \in [0,1]\} \subset \Lambda^h$. By Proposition 4.4.2, T admits a stable foliation and Y^h is union of stable leaves for T_1. Let τ^h be the first return of T_1 to Y^h. It remains to show that $\mu^h(\tau^h > n) \leq Cn^{-\alpha}$ and condition (4.1).

For the first part, note that

$$\tau^h(y,u) \leq \tau^h(y,0) \leq \left\lceil \sum_{j=0}^{\tau(y)-1} h(T^j y) \right\rceil + 1 \leq \|h\|_\infty \tau(y) + 2.$$

Therefore

$$\mu^h(\tau^h > n) \leq \mu(\|h\|_\infty \tau + 2 > n) \leq Cn^{-\alpha}.$$

For the second part, since v is Hölder, it suffices to show $T_1^{\tau^h}$ contracts stable leaves exponentially (see equation (4.4)). Let $(x,u) \in \Lambda^h$ and define the lap number $N = N(x,u,t)$ as the integer such that $u + t \in [h_N(x), h_{N+1}(x))$. Then $T_1^{\tau^h}(x,u) = T_{\tau^h}(x,u) = \big(T^N x, u + t - h_N(x)\big) = \big(T^\tau x, u + t - h_N(x)\big) = \big(Fx, u + t - h_N(x)\big)$. Let $n \geq 1$ and note that $(T_1^{\tau^h})^n(x,u) = \big(F^n x, g(x,u,n)\big)$ for some $g(x,u,n) \in [0,1]$. Therefore for any $(x',u) \in W^s(x,u)$ we have

$$d\big(T_{\tau^h}^n(x',u), T_{\tau^h}^n(x,u)\big) = d(F^n x', F^n x) + |g(x',u,n) - g(x,u,n)| \leq C\gamma^n + 1 = \tilde{C}\gamma^n,$$

and so the induced map $T_1^{\tau^h}$ contracts stable leaves exponentially as desired. □

Define $C^{\eta,m}(M) := \{w : M \to \mathbb{R} : \|w\|_{C^{\eta,m}} := \sum_{j=0}^m \|\partial_t^j w\|_{C^\eta} < \infty\}$.

Proposition 4.4.4. *Suppose* $T : \Lambda \to \Lambda$ *is modelled by a Young tower with polynomial tail* τ*, so* $\mu(\tau > n) \leq Cn^{-\beta}$ *for some* $\beta > 1$*. Assume condition* (3.14) *is* *satisfied and absence of approximate eigenfunctions holds for T_t [37, Definition 3.4]. Suppose $v \in C^{\eta}(\Lambda^h) \cap C^{0,\eta}(\Lambda^h)$ is a mean zero observable. Then,*

$$\|\mathbb{E}_0(v \circ T_1^{-n})\|_1 \leq Cn^{-\beta}.$$

Proof. [2, Corollary 8.1] guarantees an $m \geq 1$ and $C > 0$ such that

$$\left|\int v\, w \circ T_t \, d\mu^h\right| \leq C\big(\|v\|_{C^\eta} + |v|_{C^{0,\eta}}\big)\|w\|_{C^{\eta,m}} t^{-\beta},$$

for every $w \in C^{\eta,m}(\Lambda^h)$. Therefore $\left|\int v\, w \circ T_1^n \, d\mu^h\right| \leq C\|w\|_{C^{\eta,m}} n^{-\beta}$. Denote the set of continuous functions, constant on stable leaves, by $C^0(\mathcal{A}_0)$. Note that $C^0(\mathcal{A}_0) \subset C^{\eta,m}(\Lambda^h)$, which shows $\left|\int v\, w \circ T_1^n \, d\mu^h\right| \leq C\|w\|_{C^0(\mathcal{A}_0)} n^{-\beta}$ for every $w \in C^0(\mathcal{A}_0)$. Let $w' \in L^\infty(\mathcal{A}_0)$. For any $\epsilon > 0$ there exists $w \in C^0(\mathcal{A}_0)$ such that $\|w\|_{C^0(\mathcal{A}_0)} \leq \|w'\|_{L^\infty(\mathcal{A}_0)}$ and $\int |w' - w| \, d\mu < \epsilon$. Therefore

$$\begin{aligned}\left|\int v\, w' \circ T_1^n \, d\mu^h\right| &\leq \left|\int v\, w \circ T_1^n \, d\mu^h\right| + \left|\int v\,(w' - w) \circ T_1^n \, d\mu^h\right| \\ &\leq C\|w\|_{C^0(\mathcal{A}_0)} n^{-\beta} + \|v\|_\infty \|w' - w\|_1 \\ &\leq C\|w\|_{L^\infty(\mathcal{A}_0)} n^{-\beta} + \|v\|_\infty \epsilon.\end{aligned}$$

Theorem 4.0.1 now shows

$$\|\mathbb{E}_0(v \circ T_1^{-n})\|_1 \leq Cn^{-\beta}.$$

□

An Example

For $0 < e, k < 1$, let $T_{ek} : \Lambda \to \Lambda$, $\Lambda = [0,1]^2$ be the intermittent Baker's map defined in (4.3),

$$T_{ek}(x,y) = \begin{cases} (f_e(x), f_k^{-1}(y)), & (x,y) \in [0,1/2) \times [0,1] \\ (2x - 1, (y+1)/2), & (x,y) \in [1/2,1] \times [0,1] \end{cases},$$

where $f_e(x) = x(1 + 2^e x^e)$. Denote $T := T_{ek}$. Define $Y := [1/2] \times [0,1]$ and recall from the proof of Proposition 4.0.5 that T is modelled by a Young tower with tail $\tau_e := \inf\{n \geq 1 \, : \, T_{ee}^n \in Y\}$, and $\mu(\tau_e > n) \leq Cn^{-1/e}$. Therefore the iterated WIP already holds for $0 < k, e < 1/2$ by [40, Corollary 2.3]. To obtain a new example, we lift our map to a suspension flow and consider the time-one map.

Let h be a Hölder continuous and bounded roof function such that $\inf h > 0$, $\inf_Y h > 1$ and h is constant on stable leaves for T. Let $T_t : \Lambda^h \to \Lambda^h$ be the corresponding suspension flow, defined by (3.15), and assume T_t satisfies absence of approximate eigenfunctions and condition (3.14). Suppose $v \in C^\eta(\Lambda^h) \cap C^{0,\eta}(\Lambda^h)$ is a mean zero observable. Proposition 4.4.4 now shows

$$\|\mathbb{E}_0(v \circ T_1^{-n})\|_1 \leq Cn^{-1/e+1}.$$

For the remaining condition of ($\mathrm{A}_{\mathrm{strong}}$), we mimic the the argument from the proof of Proposition 4.0.5. The argument will show condition (4.1′) holds and we can apply Proposition 4.0.2. Denoting $S := T_{kk}$, let $S_t : \Lambda^h \to \Lambda^h$ be the corresponding suspension flow and $\tau_k^h(\underline{x}, u) := \inf\{n \geq 1 : S_1^n(\underline{x}, u) \in Y^h\}$. The stable foliation for S_1 is the same as the stable foliation for T_1. For the proof of Proposition 4.4.3 to carry over, we must show that $S_1^{\tau_k^h}$ contracts stable leaves exponentially and that equation (4.5) holds when we replace $T_{ek} = T$ with T_1 and T_{kk} with S_1. Exponential contraction along stable leaves for S_1 was shown in the proof of Proposition 4.4.3. To show equation (4.5) holds, it suffices to show that distances on stable leaves with respect to T_1 and S_1 agree. Let $(\underline{x}, u) \in \Lambda^h$, $(\underline{x}', u) \in W^s(\underline{x}, u)$ and let $N = N(\underline{x}, u, n)$ be the lap number, i.e. the integer N such that $u + n \in [h_N(\underline{x}), h_{N+1}(\underline{x}))$. Since h is constant on stable leaves we have $h_m(x') = h_m(x)$ for all $m \geq 1$ and $N(\underline{x}', u, n) = N(\underline{x}, u, n)$. Calculate that

$$\begin{aligned} &d\big(T_1^n(\underline{x}', u), T_1^n(\underline{x}, u)\big) = d\left[\big(T^N\underline{x}', u + n - h_N(x')\big), \big(T^N\underline{x}, u + n - h_N(x)\big)\right] \\ &\quad = d\big(T^N\underline{x}', T^N\underline{x}\big) + \big|u + n - h_N(x') - (u + n - h_N(x'))\big| \end{aligned}$$

Again, since the contraction along stable leaves does not depend on the first coordinate, we can replace $T = T_{ek}$ with $S = T_{kk}$ to obtain

$$\begin{aligned} d\big(T_1^n(\underline{x}', u), T_1^n(\underline{x}, u)\big) &= d\big(S^N\underline{x}', S^N\underline{x}\big) + \big|u + n - h_N(x') - (u + n - h_N(x))\big| \\ &= d\big(S_1^n(\underline{x}', u), S_1^n(\underline{x}, u)\big). \end{aligned}$$

Therefore equation (4.5) holds and the proof of Proposition 4.4.3 carries over. We now have

$$\|E_0(v \circ T_1^n) - v \circ T_1^n\|_1 \leq Cn^{-1/c+1}.$$

Assumption ($\mathrm{A}_{\mathrm{strong}}$) and hence the iterated WIP (Theorem 4.4.1) holds for all $0 < c < 1/3$ and $0 < e < 1/2$. This extends the previous best known result by Kelly and Melbourne [29, Theorem 5.2] significantly, which was previously known for all

$0 < c, e < 1/3$.

4.4.2 Proof of the Iterated WIP

For the proof of the iterated WIP, we will need a second coboundary decomposition

$$v = \widehat{v} + \chi_0^\infty \circ T - \chi_0^\infty, \tag{4.17}$$

where we recall $\chi_0^\infty = \sum_{n=0}^\infty \mathbb{E}_0(v \circ T^n) - v \circ T^n$. Under the assumption ($\mathrm{A}_{\text{strong}}$) we have $\chi_0^\infty \in L^2(\Lambda, \mathbb{R}^d)$. Moreover, $\widehat{v}$ is $\mathcal{A}_0$-measurable and $\widehat{v} \in L^2(\Lambda, \mathbb{R}^d)$.

To use Proposition 3.4.8 for the approximation of the stochastic integral $\int X_n \otimes dY_n \to_w \int X \otimes dY$, X_n must be an adapted process. To this end, define the processes

$$\widehat{W}_n := n^{-1/2} \sum_{j=0}^{[nt]-1} \widehat{v} \circ T^j, \qquad \widehat{W}_n^- := n^{-1/2} \sum_{j=1}^{[nt]} \widehat{v} \circ T^{-j}.$$

We see that $\widehat{W}_n^-$ is an adapted process with respect to $\{\mathcal{A}_{n,t} = T^{[nt]}\mathcal{A}_0 \,|\, 0 \leq t \leq 1\}$. Further define the processes

$$\begin{aligned}
\mathbb{X}_n^{\beta\gamma}(t) &= n^{-1} \sum_{j=0}^{[nt]-1} \sum_{i=0}^{j-1} m^\beta \circ T^i \, \widehat{v}^\gamma \circ T^j, \\
\mathbb{X}_n^{\beta\gamma,-}(t) &:= n^{-1} \sum_{j=1}^{[nt]-1} \sum_{i=j+1}^{[nt]} \widehat{v}^\beta \circ T^{-j} \, m^\gamma \circ T^{-i} = \int_0^t \widehat{W}_n^{\beta,-} \, dM_n^{\gamma,-},
\end{aligned}$$

for coordinates $\beta, \gamma = 1, \ldots, d$.

Recall that for random vectors X_j and $S_n = \sum_{j=1}^n X_j$ we have $\overline{S_n} = \max_{1 \leq \ell \leq n} |S_\ell|$. In particular,

$$\begin{aligned}
\sup_{0 \leq t \leq 1} \left| W_n^-(t) - \widehat{W}_n^-(t) \right| &= \sup_t \left| n^{-1/2} \sum_{j=1}^{[nt]} (v - \widehat{v}) \circ T^{-j} \right| \\
&= n^{-1/2} \max_{1 \leq \ell \leq n} \left| \sum_{j=1}^{\ell} (v - \widehat{v}) \circ T^{-j} \right| = n^{-1/2} \overline{(v - \widehat{v})_n^-}.
\end{aligned}$$

Proposition 4.4.5. *We have that*

$$(W_n, \mathbb{X}_n) \to_w (W, \mathbb{X})$$

in $\big(D([0,1], \mathbb{R}^d \times \mathbb{R}^{d \times d}), \|\cdot\|_\infty\big)$, *where* $\mathbb{X}^{\beta\gamma}(t) := \int_0^t W^\beta \, dW^\gamma$ *for* $t \in [0,1]$, $\beta, \gamma =$

$1, \dots, d$.

Proof. We first verify the hypothesis of Proposition 3.4.8 with $X_n = \widehat{W}_n^-$, $Y_n = M_n^-$ to obtain convergence of the backwards process $(\widehat{W}_n^-, M_n^-, \mathbb{X}_n^-) \to_w (W, W, \mathbb{X})$. As $\widehat{v}$ and m are $\mathcal{A}_0$-measurable, we see $(\widehat{W}_n^-, M_n^-)$ is a measurable process with respect to $\{\mathcal{A}_{n,t} = T^{[nt]}\mathcal{A}_0 \,|\, 0 \le t \le 1\}$. Moreover, M_n^- is a martingale by Proposition 4.3.1. The estimate $\sup_n \mathbb{E}\Big(\big[M_n^{\beta,-}, M_n^{\beta,-}\big](t)\Big) < \infty$ follows from the corresponding non-invertible calculation (equation (3.20)). It remains to show convergence of the pair $(\widehat{W}_n^-, M_n^-) \to_w (W, W)$. The martingale WIP (Proposition 3.3.5) was shown in the non-invertible case in Proposition 3.3.8, and the proof carries over to the invertible case by dropping the tilde notation. That is, $M_n^- \to_w W$, where W is a mean zero Brownian motion with covariance $\Sigma = \int m\, m^T \, d\mu$. By Proposition 4.2.5, $\Sigma = \lim_{n\to\infty} n^{-1} \int v_n \, v_n^T \, d\mu$. The continuous mapping theorem (Proposition 3.2.9) now shows $(M_n^-, M_n^-) \to_w (W, W)$. Thus $(W_n^-, M_n^-) \to_w (W, W)$ by Proposition 4.1.8. It now remains to show

$$\sup_{0\le t\le 1} \Big| W_n^-(t) - \widehat{W}_n^-(t) \Big| = n^{-1/2} \max_{1\le \ell\le n} \left| \sum_{j=1}^{\ell} (v - \widehat{v}) \circ T^{-j} \right| = n^{-1/2} \max_{\ell} \Big| \chi^+ - \chi^+ \circ T^{-\ell} \Big|$$
$$\le n^{-1/2} |\chi^+| + n^{-1/2} \max_{\ell} \big|\chi^+\big| \circ T^{-\ell} \to_p 0.$$

The first term clearly converges to zero, and the second term converges to zero by Proposition 2.2.3. Hence we have $(\widehat{W}_n^-, M_n^-) \to_w (W, W)$ and Proposition 3.4.8 now applies, so

$$(\widehat{W}_n^-, M_n^-, \mathbb{X}_n^-) \to_w (W, W, \mathbb{X}) \tag{4.18}$$

in $\big(D([0,1], \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}), \|\cdot\|_\infty\big)$.

It now remains to complete the time reversal procedure. Recall that $g(u)(t) = u(1) - u(1-t)$ and $h(u,v)(t) = u(1-t)\big(v(1) - v(1-t)\big)^T$ are continuous maps. The time reversal argument (Proposition 3.4.7) carries over to the invertible case by replacing $\tilde{W}_n^-$ with $\widehat{W}_n^-$ in the Stieltjes integral and dropping the tilde notation throughout. This gives

$$\big(W_n, \mathbb{X}_n\big) =_d \left(g(W_n^-), \Big(g(\mathbb{X}_n^-) - h(\widehat{W}_n^-, M_n^-) \Big)^T \right) + F_n,$$

where F_n is such that $\sup_{0\le t\le 1} |F_n(t)| \to_p 0$. By Proposition 3.3.12 it now suffices to show

$$\left(g(W_n^-), \Big(g(\mathbb{X}_n^-) - h(\widehat{W}_n^-, M_n^-) \Big)^T \right) \to_w Z,$$

where Z is a process with continuous sample paths and $Z =_d (W, \mathbb{X})$. It follows from (4.18) that $(W_n^-, \widehat{W}_n^-, M_n^-, \mathbb{X}_n^-) \to_w (W, W, W, \mathbb{X})$. The continuous mapping theorem (Proposition 3.2.9) now shows

$$\Big(g(W_n^-), \Big(g(\mathbb{X}_n^-) - h(\widehat{W}_n^-, M_n^-)\Big)^T\Big) \to_w \Big(g(W), (g(\mathbb{X}) - h(W,W))^T\Big).$$

Proposition 3.4.12 shows $\Big(g(W), \big(g(\mathbb{X}) - h(W,W)\big)^T\Big) =_d (W, \mathbb{X})$ and so we have shown $(W_n, \mathbb{X}_n) \to_w (W, \mathbb{X})$ as required. □

Proof of Theorem 4.4.1. For this proof we denote $\chi_0^\infty = \chi^+$, with coordinates $\chi^{\beta,+}$ for $\beta = 1, \ldots, d$. With this notation, our secondary coboundary decomposition (4.17) is $v = \widehat{v} + \chi^+ \circ T - \chi^+$.

We first show that

$$\Gamma^{\beta\gamma}(t) = \sum_{j=1}^{\infty} \int v^\beta \, v^\gamma \circ T^j \, d\mu = \int \chi^\beta \, v^\gamma - m^\beta \, \chi^{\gamma,+} \circ T \, d\mu.$$

Consider

$$\begin{aligned}
&\sum_{j=1}^{n} \int v^\beta \, v^\gamma \circ T^j \, d\mu = \sum_j \int (m^\beta + \chi^\beta \circ T - \chi^\beta) \, v^\gamma \circ T^j \, d\mu \\
&= \sum_j \int m^\beta \, (\widehat{v}^\gamma + \chi^{\gamma,+} \circ T - \chi^{\gamma,+}) \circ T^j + \int (\chi^\beta - \chi^\beta \circ T^{-n}) \, v^\gamma \, d\mu \\
&= \sum_j \int m^\beta \, \widehat{v}^\gamma \circ T^j + \int m^\beta \, (\chi^{\gamma,+} \circ T^{n+1} - \chi^{\gamma,+} \circ T) \, d\mu + \int (\chi^\beta - \chi^\beta \circ T^{-n}) \, v^\gamma \, d\mu \\
&= \sum_j \int m^\beta \, \widehat{v}^\gamma \circ T^j + \int \big(\chi^\beta \, v^\gamma - m^\beta \, \chi^{\gamma,+} \circ T\big) \, d\mu + \int m^\beta \, \chi^{\gamma,+} \circ T^{n+1} \, d\mu \\
&\qquad - \int \chi^\beta \, v^\gamma \circ T^n \, d\mu.
\end{aligned}$$

The first term is zero since for every $j \geq 1$, $\widehat{v} \circ T^j$ is $\mathcal{A}_j$-measurable and $\mathbb{E}_j m = \mathbb{E}_j \mathbb{E}_1 m = 0$. Thus $\int m^\beta \, \widehat{v}^\gamma \circ T^j \, d\mu = \int \mathbb{E}_j \big(m^\beta \, \widehat{v}^\gamma \circ T^j\big) \, d\mu = \int \mathbb{E}_j m^\beta \, \widehat{v}^\gamma \circ T^j \, d\mu = 0$. The third and fourth term converge to zero by mixing and the fact that $\int m^\beta \, d\mu = \int v^\gamma \, d\mu = 0$. Therefore $\sum_{j=1}^n \int v^\beta \, v^\gamma \circ T^j \, d\mu \to_{a.s.} \int \chi^\beta \, v^\gamma - m^\beta \, \chi^{\gamma,+} \circ T \, d\mu$.

Next, note that

$$\begin{aligned}
v^\beta \circ T^i \, v^\gamma \circ T^j &= m^\beta \circ T^i \, v^\gamma \circ T^j + (\chi^\beta \circ T - \chi^\beta) \circ T^i \, v^\gamma \circ T^j \\
&= m^\beta \circ T^i \, \widehat{v}^\gamma \circ T^j + m^\beta \circ T^i \, (\chi^{\gamma,+} \circ T - \chi^{\gamma,+}) \circ T^j + (\chi^\beta \circ T - \chi^\beta) \circ T^i \, v^\gamma \circ T^j.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\mathbb{W}_n^{\beta\gamma}(t) - \mathbb{X}_n^{\beta\gamma}(t) = n^{-1} \sum_{0\le i<j\le [nt]-1} \Big(v^\beta \circ T^i \, v^\gamma \circ T^j - m^\beta \circ T^i \, \widehat{v}^\gamma \circ T^j \Big) \\
&= n^{-1} \sum_{0\le i<j\le [nt]-1} \Big(m^\beta \circ T^i \, (\chi^{\gamma,+} \circ T - \chi^{\gamma,+}) \circ T^j + (\chi^\beta \circ T - \chi^\beta) \circ T^i \, v^\gamma \circ T^j \Big) \\
&= n^{-1} \sum_{i=0}^{[nt]-1} \sum_{j=i+1}^{[nt]-1} m^\beta \circ T^i \, (\chi^{\gamma,+} \circ T - \chi^{\gamma,+}) \circ T^j + n^{-1} \sum_{j=0}^{[nt]-1} \sum_{i=0}^{j-1} (\chi^\beta \circ T - \chi^\beta) \circ T^i \, v^\gamma \circ T^j \\
&= n^{-1} \sum_{i=0}^{[nt]-1} m^\beta \circ T^i \big(\chi^{\gamma,+} \circ T^n - \chi^{\gamma,+} \circ T^{i+1} \big) + n^{-1} \sum_{j=0}^{[nt]-1} \big(\chi^\beta \circ T^j - \chi^\beta \big) \, v^\gamma \circ T^j \\
&= n^{-1} \sum_{i=0}^{[nt]-1} \big(\chi^\beta v^\gamma - m^\beta \chi^{\gamma,+} \circ T \big) \circ T^i + n^{-1} \chi^{\gamma,+} \circ T^n \sum_{i=0}^{[nt]-1} m^\beta \circ T^i - n^{-1} \chi^\beta \sum_{i=0}^{[nt]-1} v^\gamma \circ T^i \\
&= A_n^{\beta\gamma}(t) + B_n^{\beta\gamma}(t) - C_n^{\beta\gamma}(t)
\end{aligned}$$

We see that $A_n^{\beta\gamma}(1) \to_{a.s.} \int \chi^\beta \, v^\gamma - m^\beta \, \chi^{\gamma,+} \circ T \, d\mu = \Gamma^{\beta\gamma}(1)$ by the Ergodic Theorem, and so $\sup_t |A_n^{\beta\gamma}(t) - \Gamma^{\beta\gamma}(t)| \to_p 0$ by Proposition 2.2.2. Similarly, $C_n^{\beta\gamma}(1) \to_{a.s} \chi^\beta \int v^\gamma \, d\mu = 0$ and so $\sup_t |C_n^{\beta\gamma}(t)| \to_p 0$. Finally, we see that $\sup_t |B_n^{\beta\gamma}(t)| \to_p 0$ since

$$\begin{aligned}
&\left\| \sup_t \left| n^{-1} \chi^{\gamma,+} \circ T^n \sum_{i=0}^{[nt]-1} m^\beta \circ T^i \right| \right\|_1 \le n^{-1} \|\chi^{\gamma,+}\|_2 \left\| \sup_t \left| \sum_{i=0}^{[nt]-1} m^\beta \circ T^i \right| \right\|_2 \\
&\le n^{-1} n^{1/2} \|\chi^{\gamma,+}\|_2 \|m^\beta\|_2 \to 0,
\end{aligned}$$

where we have used Doob's inequality (Proposition 3.3.9) and Proposition 4.1.4. $\square$

Bibliography

[1] J. Aaronson. *An Introduction to Infinite Ergodic Theory.* Number 50 in Math. Surveys and Monographs. Amer. Math. Soc., 1997.

[2] P. Bálint, O. Butterley, and I. Melbourne. Polynomial decay of correlations for flows, including Lorentz gas examples. *Commun. Math. Phys.*, 368:55–111, 2019.

[3] P. Bálint and I. Melbourne. Decay of correlations and invariance principles for dispersing billiards with cusps, and related planar billiard flows. *J. Stat. Phys.*, 133:435–447, 2008.

[4] M. Benedicks and L.-S. Young. Markov extensions and decay of correlations for certain Hénon maps. *Asterisque*, 261:13–56.

[5] P. Billingsley. The Lindeberg-Lévy Theorem for martingales. *Proc. Amer. Math. Soc.*, 12:788–792, 1961.

[6] P. Billingsley. *Probability and Measure.* Wiley, 1995.

[7] P. Billingsley. *Convergence of Probability Measures.* Wiley, 2nd edition, 1999.

[8] B.M. Brown. Martingale central limit theorems. *Ann. Math. Stat.*, 42(1):59–66, 1971.

[9] L. A. Bunimovich, Y. G Sinaĭ, and N. I. Chernov. Statistical properties of two-dimensional hyperbolic billiards. *Russian Math. Surveys*, 46:47–106, 1991.

[10] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne, and H. Zhang. Multiscale systems, homogenization, and rough paths. *Probability and Analysis in Interacting Particle Systems: In Honor of S.R.S. Varadhan, Berlin, August, 2016*, 283:17–42, 2019.

[11] J. Dedecker and R. Rio. On the functional central limit theorem for stationary processes. *Annales de l'I.H.P. Probabilitès et statistiques*, 36:1–34, 2000.

[12] M. Demers, I. Melbourne, and M. Nicol. Martingale approximations and anisotropic Banach spaces with an application to the time-one map of a Lorentz gas. Preprint, 2018.

[13] M. Denker and W. Phillip. Approximation by Brownian motion for Gibbs measures and flows under a function. *Ergod. Th. & Dynam. Sys.*, 4:541–552, 1984.

[14] M. D. Donsker. An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.*, 6, 1951.

[15] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley & Sons, 1985.

[16] W. Feller. *An Introduction to Probability Theory and its Applications*, volume 1. John Wiley & Sons, 1968.

[17] M. Gordin and M. Peligrad. On the functional central limit theorem via martingale approximation. *Bernoulli*, 17:424–440, 2011.

[18] M.I. Gordin. The central limit theorem for stationary processes. *Soviet. Math. Dokl.*, 10:1174–1176, 1969.

[19] G. A. Gottwald and I. Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proc. R. Soc. A.*, 469, 2013.

[20] S. Gouëzel. Central limit theorem and stable laws for intermittent maps. *Prob. Theory Relat. Fields*, 128:82–122, 2004.

[21] P. Hall and C.C. Heyde. *Martingale Limit Theory and its Application*, volume 31 of *Prob. and Math. Stat.* Academic Press, 1980.

[22] F. Hofbauer and G. Keller. Ergodic properties of invariant measures for piecewise monotonic transformations. *Math. Z.*, 180:119–140, 1982.

[23] M. Holland and I. Melbourne. Central limit theorems and invariance principles for Lorenz attractors. *J. Lond. Math. Soc.*, 76:345–364, 2007.

[24] B. R. Hunt, T. Sauer, and J. A. Yorke. Prevalence: A translation-invariant "almost every" on infinite-dimensional spaces. *Bull. Amer. Math. Soc.*, 27(2):217–238, 1992.

[25] I. A. Ibragimov. A central limit theorem for one class of dependent random variables. *Theory Prob. Appl.*, 8:83–89, 1963.

[26] A. Jakubowski, J. Mèmin, and G. Pages. Convergence en loi des suites d'intégrales stochastiques sur l'espace $\mathbb{D}^1$ de Skorokhod. *Prob. Theory Relat. Fields*, 81:111–137, 1989.

[27] O. Kallenberg. *Foundations of Modern Probability.* Springer-Verlag New York, 2nd edition, 2001.

[28] G. Keller. Un théorème de la limite centrale pour une classe de transformations monotones par morceaux. *C. R. Acad. Sci. Paris Sér. A-B*, 291:155–158, 1980.

[29] D. Kelly and I. Melbourne. Smooth approximation of stochastic differential equations. *Ann. Prob.*, 44:479–520, 2016.

[30] D. Kelly and I. Melbourne. Deterministic homogenization for fast-slow systems with chaotic noise. *J. Funct. Anal.*, 272:4063–4102, 2017.

[31] T.G. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Prob.*, 19:1035–1070, 1991.

[32] C. Liverani. Central limit theorem for deterministic systems. volume 362 of *Pitman Research Notes in Mathematics Series*, 1996.

[33] C. Liverani, B. Saussol, and S. Vaienti. A probabilistic approach to intermittency. *Ergod. Th. & Dynam. Sys.*, 19:671–685, 1999.

[34] M. Maxwell and M. Woodroofe. Central limit theorems for additive functionals of Markov chains. *Ann. Prob.*, 28:713–724, 2000.

[35] D.L. McLeish. Dependent central limit theorems and invariance principles. *Ann. Prob.*, 2:620–628, 1974.

[36] I. Melbourne. Large and moderate deviations for slowly mixing dynamical systems. *Proc. Amer. Math. Soc.*, 137(5):1735–1741, 2009.

[37] I. Melbourne. Superpolynomial and polynomial mixing for semiflows and flows. *Nonlinearity*, 31(10):R268–R316, 2018.

[38] I. Melbourne and A. Stuart. A note on diffusion limits of chaotic skew product flows. *Nonlinearity*, 24:1361–1367, 2011.

[39] I. Melbourne and A. Török. Statistical limit theorems for suspension flows. *Israel J. Math.*, 144:191–209, 2004.

[40] I. Melbourne and P. Varandas. A note on statistical properties for nonuniformly hyperbolic dynamical systems. *Stoch. Dyn*, 16, 2016.

[41] I. Melbourne and R. Zweimüller. Weak convergence to stable Lévy processes for nonuniformly hyperbolic dynamical systems. *Annales de l'I.H.P. Probabilitès et statistiques*, pages 545–556, 2015.

[42] F. Merlevéde, M. Peligrad, and S. Utev. Recent advances in invariance principles for stationary sequences. *Probability Surveys*, 3:1–36, 2006.

[43] S. V. Nagaev. Some limit theorems for stationary Markov chains. *Theory Prob. Appl.*, 2:378–409, 1957.

[44] K. Petersen. *Ergodic Theory.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1983.

[45] M. Pollicott and R. Sharp. Invariance principles for interval maps with an indifferent fixed point. *Commun. Math. Phys.*, 229:337–346, 2002.

[46] P. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Stochastic modelling and applied probability.* Springer-Verlag Berlin, 2005.

[47] M. Ratner. The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature. *Israel J. Math.*, 16:181–197, 1973.

[48] E. Rio. *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants*, volume 31 of *Mathèmatiques et Applications.* Springer-Verlag Berlin, 2000.

[49] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proc. Sixth Berkeley Symp. on Math. Stat. and Prob., Vol. 2*, pages 583–602. University of California Press, 1972.

[50] M. Tyran-Kamiǹska. An invariance principle for maps with polynomial decay of correlations. *Commun. Math. Phys.*, 260:1–15, 2005.

[51] W. Whitt. Proofs of the martingale FCLT. *Prob. Surveys*, 4:268–302, 2007.

[52] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. Math.*, 147:585–650, 1998.

[53] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.*, 110:153–188, 1999.

[54] R. Zweimüller. Mixing limit theorems for ergodic transformations. *J. Theoret. Prob.*, 20:1059–1071, 2007.